Markerless 3D Head Tracking for Motion Correction in High Resolution PET Brain Imaging

Olesen, Oline Vinter

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

REGION SIEMENS

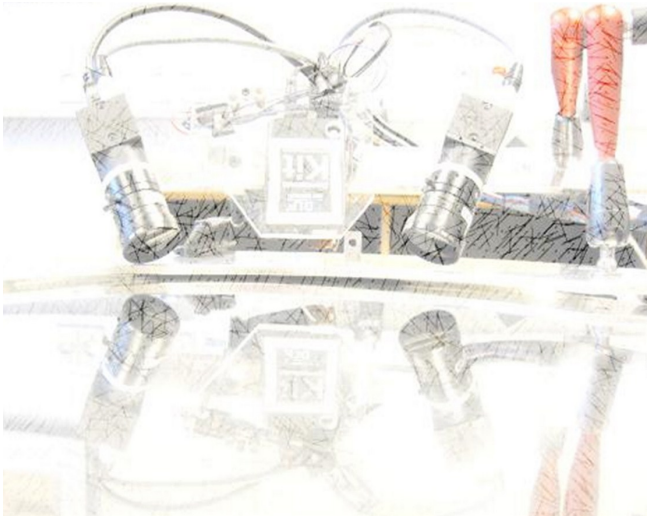
Rigshospitalet

DTU

FACULTY OF HEALTH SCIENCES

UNIVERSITY OF COPENHAGEN

Markerless 3D Head Tracking for Motion Correction in High Resolution PET Brain Imaging



Oline Vinter Olesen

Copenhagen 2011
IMM-PHD-2011-266

University
Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

Hospital
Department of Clinical Physiology, Nuclear Medicine & PET
Rigshospitalet, University of Copenhagen
Blegdamsvej 9, 2100 Cohenhagen, Denmark

Company
Siemens A/S
Healthcare Sector
Borupvang 3, 2750 Ballerup, Denmark

Printed by IMM, December 2011

IMM-PHD: ISSN 0909-3192, ISBN 9788764309300

Summary

This thesis concerns application specific 3D head tracking. The purpose is to improve motion correction in position emission tomography (PET) brain imaging through development of markerless tracking. Currently, motion correction strategies are based on either the PET data itself or tracking devices relying on markers. Data-driven motion correction is problematic due to the physiological dynamics. Marker-based tracking is potentially unreliable, and it is extremely hard to validate when the tracking information is correct. The motion estimation is essential for proper motion correction of the PET images. Incorrect motion correction can in the worst cases result in wrong diagnosis or treatment.

The evolution of a markerless custom-made structured light 3D surface tracking system is presented. The system is targeted at state-of-the-art high resolution dedicated brain PET scanners with a resolution of a few millimeters. State-of-the-art hardware and software solutions are integrated into an operational device. This novel system is tested against a commercial tracking system popular in PET brain imaging. Testing and demonstrations are carried out in clinical settings.

A compact markerless tracking system was developed with an accuracy sufficient for PET imaging (< 0.1 degrees and < 0.3 mm). Furthermore, the first non-visible structured light system using Pico DLP technology was used. In a proof-of-principle study with two human PET scans, the system was demonstrated to improve PET image quality significantly. The results were similar to motion correction using an integrated commercial marker-based system. Furthermore, phantom studies were performed supporting the system's abilities for PET motion correction.

Resumé

Denne afhandling omhandler applikationsspecifik 3D hoved tracking. Formålet er at forbedre bevægelseskorrektion i position emissions tomografi (PET) hjernescanning. Det er opnået gennem udvikling af et tracking system uden brug af markører. Nuværende bevægelseskorrektions-strategier er baseret på enten: selve PET-dataene eller eksterne tracking-systemer, som er afhængig af markører. Data-dreven bevægelseskorrektion er problematisk på grund af den fysiologiske dynamik. Markør-baseret tracking er potentielt upålidelig. Derudover er det er svært at validere, hvornår trackingen eventuelt har fejlet. Nøjagtig bestemmelsen af hovedets position er essentiel for korrekt bevægelseskorrektion af PET billederne. Forkert bevægelseskorrektion kan i værste tilfælde resultere i fejlagtig diagnose eller behandling.

De forskellige stadier i udviklingen af et markørløst struktureret lys 3D overflade tracking-system er præsenteret. Systemet er rettet mod en *state-of-the-art* dedikeret hjerne PET-scanner med en rummelig opløsning på kun få millimeter. *State-of-the-art* hardware og software løsninger er integreret i en operationel enhed. Dette nye system er testet mod et kommercielt tracking-system, som er populært i PET hjernescanning. Test og demonstrationer er udført i kliniske omgivelser.

Et kompakt tracking-system uden brug af markører blev udviklet med en nøjagtighed tilstrækkelig til PET-scanning ($< 0{,}1$ grader og $< 0{,}3$ mm). Dette system var det første struktureret lys system ved brug af ikke-synligt lys, som anvender Pico DLP-teknologi. Det blev vist at systemet forbedrede PET billedkvaliteten væsentligt gennem et *proof-of-principle* studie inkluderende to PET-scanninger af en forsøgsperson, samt flere fantom-studier. Resultaterne var sammenlignelige med bevægelseskorrektionen ved hjælp af et integreret kommercielt markørbaseret system.

Preface

This thesis was prepared at: 1) the section for Image Analysis and Computer Graphics at Informatics Mathematical Modelling, the Technical University of Denmark (DTU), 2) the Department of Clinical Physiology, Nuclear Medicine and PET , Rigshospitalet (RH), University of Copenhagen (KU), and 3) the Healthcare Sector, Siemens A/S. Furthermore was large part of the work done at the Yale PET Center, Yale University, New Haven, USA. The thesis is a part of the requirements for the Ph.D. degree in engineering.

It is an industrial PhD project as a results of cooperation with The Danish Agency for Science, Technology and Innovation "Forsknings- og Innovationsstyrelsen". Research funding was provided by the Danish Agency for Science, Technology and Innovation and Siemens A/S.

The subject of the thesis is improving tracking of head motion during medical image acquisitions. The thesis consist of a selection of research papers written during the period 2008–2011.

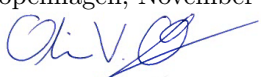
Supervisors

Professor, PhD, Rasmus Larsen, DTU
Head of Depart., Professor, Dr.Med, Liselotte Højgaard, RH and KU
Country Sector Lead, M.Sc.EE., Bjarne Roed, Siemens
Associate professor, PhD, Rasmus R. Paulsen, DTU

Assessment Committee

Professor, Dr.Sc, PhD, Søren M. Bentzen,
Department of Human Oncology, University of Wisconsin School of Medicine and Public Health, Madison, USA
Section Lead, PhD, Mauricio Reyes,
Inst. for Surgical Technology & Biomechanics, University of Bern, Switzerland
Professor, Knut Conradsen,
DTU-Informatics, Technical University of Denmark, Denmark

Copenhagen, November 2011



Oline Vinter Olesen

Papers included in the thesis

[A] Oline V. Olesen, Claus Svarer, Merence Sibomana, Sune H. Keller, Søren Holm, Jørgen A. Jensen, Flemming Andersen, and Liselotte Højgaard. A Moveable Phantom Design for Quantitative Evaluation of Motion Correction Studies on High Resolution PET Scanners. *IEEE Transactions on Nuclear Science*, vol. 57, p. 1116-1124, 2010. ([116])

[B] Oline V. Olesen, Morten R. Jørgensen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Structured Light 3D Tracking System for Measuring Motions in PET Brain Imaging. *Proceedings of SPIE, The International Society for Optical Engineering*,vol. 7625, p. 76250X-11, 2010. ([108])

[C] Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Motion Tracking for Medical Imaging: A Non-Visible Structured Light Tracking Approach. *IEEE Transactions on Medical Imaging*, 2011. ([112])

[D] Oline V. Olesen, Rasmus R. Paulsen, Rasmus R. Jensen, Sune H. Keller, Merence Sibomana, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. 3D Surface Realignment Tracking for Medical Imaging: A Phantom Study with PET Motion Correction. *Image-based Geometric Modeling and Mesh Generation by Springer, E. Yongjie Zhang.* Accepted for publication ([113])

[E] Oline V. Olesen, Jenna M. Sullivan, Tim Mulnix, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, Richard E. Carson, Evan D. Morris, and Rasmus Larsen. List-mode PET motion correction using markerless head tracking: proof-of-concept in human studies. Submitted for journal publication.

Additional papers by the author

1 Sune H. Keller, Merence Sibomana, Oline V. Olesen, Claus Svarer, Søren Holm, Flemming Andersen, and Liselotte Højgaard. Methods for motion correction evaluation using FDG human brain scans on a high resolution PET scanner. *Journal of Nuclear Medicine* . Accepted for publication ([83])

2 Oline V. Olesen, Rasmus R. Paulsen, Sune H. Keller, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Comparison of external motion tracking systems for PET list-mode reconstruction. *IEEE Nuclear Science Symposium Conference Record, 2011* . ([114])

3 Jakob Wilm, Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Real Time Surface Registration for PET Motion Tracking. *Lecture Notes in Computer Science, Springer, vol. 6688, p. 166-175, 2011* . ([166])

4 Martin Kjer, Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Geometric calibration between PET scanner and structured light scanner. *Proc. of the MICCAI workshop on Mesh Processing in Medical Image Analysis, 2011* . ([84])

5 Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. Motion tracking in narrow spaces: a structured light approach. *Lecture Notes in Computer Science, Springer, vol. 6363, p. 253-260, 2010.* ([111])

6 Oline V. Olesen, Sune H. Keller, Merence Sibomana, Rasmus Larsen, Bjarne Roed, and Liselotte Højgaard. Automatic Thresholding for Frame-Repositioning Using External Tracking in PET Brain Imaging. *IEEE Nuclear Science Symposium Conference Record, p. 2669-2675, 2010.*([109])

7 Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen. External motion tracking for brain imaging: structured light tracking with invisible light. *IEEE Nuclear Science Symposium Conference Record, p. 2735-2737, 2010.* ([110])

8 Oline V. Olesen, Merence Sibomana, Sune H. Keller, Flemming Andersen, Jørgen A. Jensen, Søren Holm, Claus Svarer, and Liselotte Højgaard. Spatial resolution of the HRRT PET scanner using 3D-OSEM PSF reconstruction. *IEEE Nuclear Science Symposium Conference Record, p. 3789-3790, 2009* . ([115])

9 Merence Sibomana, Sune H. Keller, Claus Svarer, Oline V. Olesen, Flemming Andersen, Søren Holm, and Liselotte Højgaard. New attenuation

correction for the HRRT using transmission scatter correction and total variation regularization. *IEEE Nuclear Science Symposium Conference Record, p. 3284-3286, 2009.* ([147])

Acknowledgements

I want to express my thank you and warmest regards to supervisors Professor Rasmus Larsen, Professor Liselotte Højgaard, and Country Lead Bjarne Roed for positive and constructive corporation in making this project successful. It has been an honor to benefit from your solid experience and knowledge. I am very proud of the great confidence they have shown in me. I would like to thank them for their engagement in the project and for their ability to motivate - as well as challenge me.

I am deeply grateful for all the help and supervision Rasmus Larsen has given me over the years - both professionally and morally. His door was always open when I needed some advice. Thank you! It is important to me to extend my deep gratitude to Liselotte Højgaard for always wanting the best for me and for giving me the opportunity to perform the project at the Department of Clinical Physiology, Nuclear Medicine and PET at Rigshospitalet - the Copenhagen University Hospital. This is also due to the Birthe and John Meyer Foundation, which has sponsored the HRRT scanner. Furthermore, special thank you to Bjarne Roed for raising the funding from Siemens and for introducing me to Siemens internationally (Knoxville, USA and Erlangen, Germany) and providing me with world-wide experience through many journeys.

I would also like to thank Associate Professor Rasmus R. Paulsen who has been very helpful with all kinds of issues especially related to surface reconstruction. Rasmus has also very supportingly taking part of the supervision. Rasmus as well as the persons I sharing the office with: Rasmus R. Jensen and Jens Fagertun, have been very supporting and helpful and kept a good spirit in the office.

The Department of Clinical Physiology, Nuclear Medicine and PET at Rigshospitalet has supported the project in every manner, making the work possible

in a pleasant atmosphere. I would like to thank Sune H. Keller PhD and medical imaging consultant Merence Sibomana for their helpfulness and positive corporations.

The Yale PET center, New Haven, USA led by Professor Richard E. Carson kindly hosted me during a four-month stay. I would like to express a great thank you for allowing me to test my equipment at the HRRT setting at Yale. I enjoyed the time I spent at Yale, and I appreciate all the kind people I met there. I am especially in debt to Tim Mulnix, Engineer, who kindly and effectively helped me integrate my system with the HRRT scanner at the Yale PET center. I would also like to thank Associate Professor Evan D. Morris and Jenna M. Sullivan, graduate student for their kindness and for giving me the opportunity of being a part for their PET studies.

I would also like to mention Michael Rasmussen, who provided software for running the developed system; work Michael mainly did during his Bachelor project. Jakob Wilm and Martin Kjer also provided good input to this project during their Bachelor and special course projects.

I owe a great debt to PhD candidate Otto Nielsen MSc in Physics, for kindly performing optical measurements of the system. Otto made sure that the system characterization regarding eye safety was professionally carried out.

Jens Olsen and the mechanical workshop at Rigshospitalet have made high quality devices at very short notice. Thank you!

Many volunteers have participated in experiments related to this project. These volunteers have given a fundamental understanding of the tracking system and movements during PET brain imaging. I appreciate their readiness and I would like to say a warm thank you.

Finally, I would like to express my sincerest thanks to my family and Marie Riisgaard who have lovingly supported me throughout the project and for understanding my high commitment and engagement to the project.

Abbreviations

AIR	Automated image registration
ANOVA	Analysis of variance between groups
API	Application programmers interface
As	Arsenic
BGO	Bismuth germanate
CCDs	Charge coupled devices
CT	Computed tomography
Cu	Copper
DMD	Digital micromirror device
DLP	Digital light processing
DTU	Technical University of Denmark
ECG	Electrocardiogram
EEG	Electro encephalogram
EM	Emission
EVMS	Enterprise volume management system
F	Fluor
FBP	Filtered-back-projection
FDG	Fluoro-deoxy-glucose
FOV	Field-of-view
Ge	Germanium
GFM	GFMesstechnik GmbH
GPS	Global positioning system
HRRT	High resolution research tomograph
HVGA	Half-size video graphics array
ICP	Iterative closest point
IR	Iterative reconstruction
LED	Light emitting diode

LOR	Line-of-response
LSO	Lutetium oxyorthosilicate (Lu2SiO5)
LYSO	Cerium doped lutetium yttrium orthosilicate
MAF	Multiple acquisition frames
MC	Motion correction
MI	Mutual information
MLEM	Maximum-likelihood expectation maximization
MOLAR	The motion-compensation OSEM list-mode algorithm for resolution-recoveryreconstruction
MRI	Magnetic resonance imaging
MT	Motion tracking
μ-map	Attenuation map
NaI	Natrium iodide
NaI(Ti)	Thallium doped sodium iodide
NDI	Northern Digital Inc.
NECR	Noise effective count rate
NIH	National Institutes of Health
NIR	Near infrared
NIRED	Near infrared emitting diode
NTP	Network time protocol
O	Oxygen
OSEM	Ordered subset expectation maximization
PCA	Principal component analysis
PET	Positron emission tomography
PMT	Photomultiplier tubes
PGR	Point Grey Research
POL	The Polaris Vicra system
PSF	Point spread function
PSI	Phase-shift interferometry
QC	Quality control
RGB	Red green and blue
RMS	Root-mean-square
ROI	Region of interest
SD	Standard deviation
SIFT	Scale-invariant features
SL	Structured light
SNR	Signal to noise ratio
SPECT	Single photon emission computed tomography
SVD	Singular value decomposition
TAC	Time-activity-curve
TCL	The Tracoline system
TOF	Time-of-flight
TX	Transmission
XC	Cross-correlation

Contents

Chapter 1

Introduction

Position emission tomography (PET) is a medical imaging modality with increasing importance in clinical and research settings. In the last decades PET methodology has improved considerably due to comprehensive developments in radiochemistry and advanced scanner technology. This progress has made PET a powerful functional image modality.

PET has been explored as a source for imaging biomarkers and has become a superior tool in oncology for diagnosis, treatment planning, and follow-up. PET is also widely used for studies of pathophysiology and research. In neurology PET imaging enables in vivo visualization of cerebral blood flow, brain metabolism, and neuroreceptor binding. PET acquisition times range from 20 minutes to several hours. Because of the prolonged acquisition times, patient motion during scanning is inevitable. Patient motion during scanning will result in blurred and erroneous PET images when not corrected for. Head motion during the brain PET imaging acquisition degrades the PET image and destroys dynamic PET analysis. In worst case scenarios it can lead to fatally wrong diagnoses.

Motion correction (MC) in PET brain imaging has been extensively investigated. In general, suggested methods assume precise registered motion of the head during the scan. The Polaris Vicra system from Northern Digital Inc. (NDI, Waterloo, Canada) is a popular choice to record head motion during PET scanning. However, the problem with this system and other proposed systems is that they rely on markers attached to the head. In practice it is difficult to obtain a stable stiff coupling between the marker(s) and the patient head. Therefore this attachment is a major problem leading to incorrectly registered

motion of the head and thus erroneous motion-compensated PET reconstruction.

Motion-induced image degeneration increases with improved scanner resolution. Thus head motion is especially harmful for high resolution scanners. The dedicated PET brain scanner the high resolution research tomograph (HRRT) from Siemens has a resolution down to 1.4 mm. The subject of this PhD project is to improve 3D head tracking for MC in PET brain imaging. A new system was developed for the HRRT PET scanner and prof-of-principal subsequently obtained.

1.1 Objectives

The overall objective of this thesis was to improve motion correction strategies in medical imaging in general through the development of markerless surface tracking. The concept was to demonstrate the proof-of-principle on the HRRT dedicated brain PET scanner with potential adaptation for other scanners and imaging modalities.

The aims of this PhD project were, 1) evaluate the need for an alternative to existing state-of-the-art tracking systems in PET brain imaging; 2) design and construct a markerless 3D head tracking system targeted for the HRRT brain PET scanner with a challenging narrow bore hole geometry; 3) validate the tracking performance of the developed system against a commercial tracking system; and 4) integrate the developed system with a PET scanner and evaluate its ability for motion correction in PET brain imaging.

1.2 Scientific contributions and thesis overview

The scientific contributions of this PhD project are collected in the following five main papers upon which the thesis is written.

A O.V. Olesen, C. Svarer, M. Sibomana, S.H. Keller, S. Holm, J.A. Jensen, F. Andersen, and L. Højgaard. A Moveable Phantom Design for Quantitative Evaluation of Motion Correction Studies on High Resolution PET Scanners. *IEEE Transactions on Nuclear Science*, 2010.

B O.V. Olesen, M.R. Jørgensen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Structured Light 3D Tracking System for Measuring Motions

in PET Brain Imaging. *Proceedings of SPIE, The International Society for Optical Engineering*, 2010.

C O.V. Olesen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Motion Tracking for Medical Imaging: A Non-Visible Structured Light Tracking Approach. *IEEE Transactions on Medical Imaging*, 2011.

D O.V. Olesen, R.R. Paulsen, R.R. Jensen, S.H. Keller, M. Sibomana, L. Højgaard, B. Roed, and R. Larsen. 3D Surface Realignment Tracking for Medical Imaging: A Phantom Study with PET Motion Correction. *Image-based Geometric Modeling and Mesh Generation by Springer, E. Yongjie Zhang.*

E O.V. Olesen, J.M. Sullivan, T. Mulnix, R.R. Paulsen, L. Højgaard, B. Roed, R.E. Carson, E.D. Morris, and R. Larsen. List-mode PET motion correction using markerless head tracking: proof-of-concept in human studies.

Paper A, paper B, paper C are published, paper D has been accepted for publication in a Springer book chapter, and paper E has been submitted for journal publication. The original paper forms are presented in Appendices A–E and a list is found on page vii including additional work during the PhD study period.

Paper A describes how the commercial tracking system, Polaris Vicra from NDI, was implemented on the HRRT PET scanner at Rigshospitalet and used with the multiple acquisition frames (MAF) motion correction. The novelty of this paper was a motion-controlled PET brain phantom, developed to evaluate MC strategies in PET imaging. It was used to investigate the sources of errors related to the integrated setup of the Polaris system and the HRRT PET scanner. The study showed that the MC based on the Polaris system is accurate on phantoms where the Polaris registration tool is fastened to the phantom by metal screws. Furthermore, the study showed that the accuracy of the tracking system has a significant impact on the MC result compared to the other investigated sources of errors, including the accuracy of the geometric calibration between the tracking system and the PET scanner.

We concluded that the uncertainty of the registration tool attachment to the heads of patients renders the system inadequate for head-pose estimation for human subjects in a clinical environment. This was supported by practical experience on human PET scans with the Polaris tracking system.

In paper B we demonstrate a new markerless tracking approach based on multiple structured light (SL) surface scans. The surface scanning is based on input from two cameras and on the state-of-the-art Pico DLP projector technology.

The tracking principle was presented in [111] and paper C. Here it is shown how the SL system was tested against the Polaris system with comparable results. The SL system was modified to use non-visible light to avoid discomfort to the patient, as is also presented in paper C.

In order to apply the developed system with MC in PET brain imaging, a certain frame rate of image capture was required. The system was further improved to handle on-the-fly image capture with post tracking estimation, now named Tracoline. In paper D the Tracoline system was demonstrated with the HRRT PET scanner on a custom phantom design. The concept was similar to paper A, substituting the commercially available Polaris system with the Tracoline. An additional scientific contribution was a novel method for geometric calibration between the PET scanner and the Tracoline system.

The main question is: can the Tracoline system be used for precise human head tracking during PET acquisition with improved PET results? This is shown in paper E, based on an advanced list-mode MC reconstruction algorithm. Two human scans and one phantom scan with simultaneous Tracoline and Polaris trackings are included in this study. Data were acquired during a research stay at the Yale PET Center with Professor R. E. Carson and his group, who routinely use the list-mode event-by-event MC reconstruction MOLAR.

Thesis overview: The thesis includes an initial background part (Chapters 1–8) and a scientific contribution part (Appendices A–F). The background chapters introduce the reader to the main topics and methods in relation to motion correction in PET brain imaging, and to the basis of stereo vision for head tracking. Chapter 2 addresses motion in PET imaging and the related problems. Chapter 3 gives a brief introduction to PET and the special HRRT PET scanner. Chapter 4 contains a short review of MC in PET brain imaging, including more details of the MC methods applied in the later paper Appendices. An overview of state-of-the-art motion tracking methods for medical imaging, in particular PET brain imaging, is given in Chapter 5. Chapter 6 outlines alternative tracking methodologies and introduces the methods applied. The final background Chapter familiarizes the reader with the Tracoline system.

A thesis conclusion is found in Chapter 8, before the five Appendices presenting the scientific papers introduced above. Two final appendices concern an important safety aspect related to irradiation of human eyes. Appendix F is a technical report elaborating on the optical system characteristics and Appendix G is a system certification from an external company.

Chapter 2

Motion Induced Image Degeneration

There is a possibility of serious error on a PET brain image, if no correction is made for a patient's head movements. Head movements are a major cause of degradation of image quality for some scans, especially during long acquisition time, which can take up to hours. Head movements can also cause fatal distortions of a functional analysis due to artifacts or wrongly detected signals in the region-of-interest (ROI). Events, which have been registered outside the ROI might have occurred inside it, or vice versa.

A patient's movements, even if only a few millimeters, have a blurring effect on PET brain images on a regular, 0.5–1 cm spatial resolution, PET scanner [58]. The blurring effect increases with the acquisition time and the magnitude of movements. The most important consideration here is that with increasing scanner resolution this problem becomes even more serious. Thus a patient's movements end up counteracting the technological advances of a high resolution scanner. Therefore, motion correction must be done, if there is to be a benefit from the modern technology represented by the HRRT PET scanner.

In Fig. 2.1 the problem arising due to motion during high resolution PET brain scanning is illustrated. The images show a comparison of the image performance between a conventional PET scanner (bottom) and the HRRT scanner (top), with both scans performed on the same day on the same patient consecutively of the two types of scanners at Rigshospitalet. The image shows the same three slices of the same brain.

On the HRRT images it is possible to distinguish between gray and white matter in contrast to the images from the conventional PET scanner. If motions occur during the data acquisition, the resulting images get blurred and HRRT images end up looking like the images from a low resolution scanner. Motion correction or complete fixation of the head is necessary to utilize the high resolution of the HRRT PET scanner.

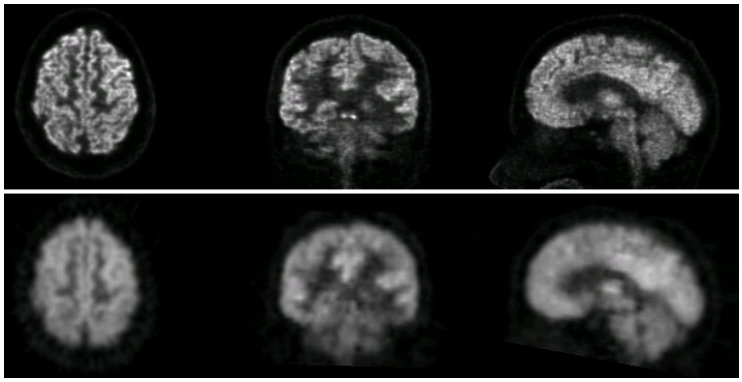


Figure 2.1: Comparison of the HRRT PET brain scanner (Top) and the GE Advance scanner (bottom). The images show three different slices of the same brain of static ^{18}F-fluoro-deoxy-glucose (FDG) scans.

Motion correction for the HRRT PET scanner has been a major issue since the first installation of the current version in 2002 [1] [22, 129, 132]. Not only has the improved scanner resolution increased the demand for motion correction, but the protocols of the PET studies have also become increasingly complex and the demand for motion correction has increased. Electro encephalograms (EEG) are often monitored during neurological PET imaging for diagnosis of epilepsy. Studies of patients with neurological or physiological diseases such as Parkinson and Alzheimer's disease studies [62, 151] are increasing. Patients often have direct involuntary motion because of their disease or due to pain and discomfort during the long scan acquisition. Furthermore, the scan subjects are often asked to perform tasks during a PET study, e.g. motor response (visually or by mind), or even smoking cigarettes during the PET scan as in paper E. This means that the scanner subjects are moving to a greater extent during the PET acquisition.

Another recently reported motivation for motion correction is scatter estimation error due to misalignment between transmission (TX) and emission scans (EM), causing wrong quantification of the PET events [7]. This study is a simulation based on a patient case experienced in an HRRT PET study with ^{11}C-verapamil

[1] Two prototypes were installed in Cologne (1999) and Amsterdam (2001).

from Amsterdam. The study showed false reduced uptake in some brain regions and the artifacts increase with the motion.

Motion correction is crucial for dynamic PET imaging where the radiotracer distribution is tracked over time. Motion destroys the true radiotracer distribution, leading to incorrect ROI analysis. Figure 2.2 shows two examples of time-activity-curves (TACs) from ^{11}C-raclopride scans presented in paper E. The TACs are for the small putamen ROIs of the brain. The blue curves represent results for a reconstruction without MC, whereas the red and green curves represent results from motion-corrected reconstructions. Clearly, ROI analysis based on the non-motion-corrected reconstruction causes an incorrect TAC curve. The observed motions of the inspected ROIs were in these cases up to 2 cm. The latter case is mainly caused by motion between the TX and EM scan of approximately 1.5–2.5 cm.

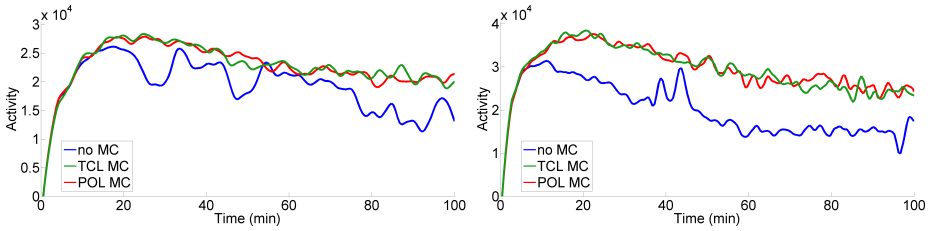


Figure 2.2: Time-activity curves for the left putamen from two ^{11}C-raclopride HRRT scans [paper E]. The scans were reconstructed with three reconstruction methods: without MC (no MC), with MC based on motion information registered with the developed Tracoline system (TCL MC), and, with MC based on motion information from the Polaris system (POL MC).

Time consuming head restraint methods have been suggested using a thermoplastic mask known from radiotherapy [67], a band-strap, and a vacuum bag [109]. However these do not eliminate the motion problem [58, 140]. Moreover, head restraints are inconvenient and uncomfortable for many patients, and they are limiting for some research studies e.g. the smoking study in paper E. In a study in Turku, Finland, 49 patients were scanned using the thermoplastic mask and the position of the head was tracked with an external tracking device reporting significant motion despite head restraints [2].

[2] Jussi Hirvonen, Turku, Presentation at the HRRT user meeting in Turku may 26, 2008.

2.1 Head and brain motions

There are pre-dominantly three kinds of head motion seen during PET scanning; long drift motion, frequent short motions around a mean position and occasional quick movements. Different amplitudes of the three motion patterns are usually represented in all scans. Long drift motion occurs when the patients are relaxing after being placed in the scanner and also if they fall asleep during scanning. Long drift motions are in general rotational motions sidewise or nodding. Frequent short motions around a mean position are related to periodic motions such as respiratory motion and neurotic motion patterns. Occasional quick movements especially occur when the patient is distracted by noise, thoughts or neurotically related incidents. A typical cause is interaction with the personnel, such as if the technologist is talking to the patient. It can also be caused by motion of the legs that mainly results in transaxial translation.

The blurring effect of the motion depends on the type of motion and the amplitude of the motion with respect to the ROI and the time period of investigation. A short peak motion during a long time frame of 20 min is much less significant than a short peak motion during a few-second frame. In contrast a long drift motion of several minutes is highly degenerating for a long time frame, while less crucial for a short time frame. The effect of the motion is also dependent on the ROI. The sensitivity towards motion increases with decreasing size of the ROI. Furthermore, the motion effect is influenced by the location of the ROI, particularly for rotary motions. The effect of rotary motion increases with the distance to the axis of rotation. Rotary motions are often around a point of contact between the patient head and the couch. This means ROIs peripheral to this point, such as the frontal lope, degenerate to a greater extent than ROIs in the back of the head. Obviously, the effect of translatory motion is uniformly distributed over the whole brain.

The magnitude of the motion is up to 2–3 cm and it has been shown that significant motion between the TX and EM scan for the HRRT PET scanner often occurs. The mentioned motion patterns are based on observations from registered motion with an external tracking system during more than 50 scans on the HRRT PET scanner at Rigshospitalet and are supported in the literature [34, 45, 65, 109].

Our aim is to correct for the movements of the brain. Therefore it is the movements of the brain that should be registered. These can be very difficult to observe directly using optical devices. Therefore, we consider the brain to be fixed to the skull as a rigid body, unchanging in shape and structure during scanning.

Non-rigid motion of the brain, i.e. brain shift, is a significant problem in intracranial surgery. The cause of brain shift during surgery has been investigated and can be in the order of centimeters [105,136]. The brain shift is mainly associated with gravity, loss of fluid and deformation of any tumors. It is in general assumed that brain shift is limited during a PET scan with no external intervention on the brain [45, 65, 102, 122]. However, both circulatory and respiratory systems cause motions within the brain, although they are considered to be smaller than the scanner resolution.

Chapter 3

Positron Emission Tomography and the High Resolution Research Tomograph

This chapter is an introduction to nuclear PET imaging and includes physics related to PET and the basics of a PET scanner. It presupposes some fundamental knowledge of quantum and radiation physics. The advanced HRRT PET scanner is the foundation for the contributed studies and it is briefly described at the end of this chapter.

3.1 Conventional PET

The basis of conventional PET is, as the name suggests, the emission of positrons from the nucleus of radioactive atoms. The existence of the positron was postulated in 1928 by Paul AM Dirac and the positron was proved to exist in 1932 by Carl Anderson. Both of them received the Nobel Prize in physics for their discoveries. The first positron scanner was introduced in the 1950s. It was a brain tumor "scanner" with only two detectors at the Massachusetts General Hospital in Boston. The positron-emitting isotope used was ^{13}As [11, p. 13].

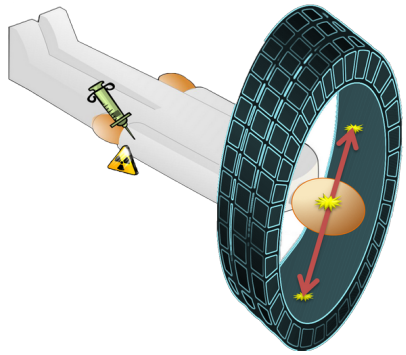


Figure 3.1: Illustration of a brain PET scan. The patient is injected with a positron-emitting tracer. After annihilation, two photons are emitted and detected in a ring of multiple small detectors.

In the 1960s a NaI multicrystal position detector, built as a ring, was developed at Brookhaven National Laboratory with a resulting spatial image resolution of approximately 2 cm. The primary isotope was ^{14}Cu [163]. Many isotopes and detectors have been developed since, but the ring structure is still a basic part of the design of the PET scanners today.

A PET scanner deviates from computed tomography (CT) imaging by detecting events inside the patients and not detecting transmissions through the patients. A patient is injected with a radiotracer. The radiotracer consists of an organic substitute, e.g. a glucose molecule, which is labeled by a positron-emitting isotope, also referred to as a radionuclide. In oncology ^{18}F-fluoro-deoxy-glucose (FDG) is the most often used radiotracer. A positron is emitted from the radionuclide and annihilates with a nearby electron, creating two photons. The two photons are emitted in opposite directions at an angle of 180^o, which the PET scanner detects in coincidence (Fig. 3.1). Coincidence means that the two photons reach two detectors within a time window of 4–12 ns and the event is registered as a true decay event.

A line between two detectors in coincidence, along which an annihilation has occurred, is called a line-of-response (LOR). The location of a single annihilation event along a LOR is unknown. However, when the acquisition of a large number of coincidence events are summed for each LOR and sorted into parallel projections, an image reflecting the tracer distribution inside the patient can be obtained [143, p. 14–6].

Depending on the biochemical structure of the radiotracer, it accumulates in

certain cells inside the patient. Cancer cells have a high glucose uptake, thus a glucose-like tracer primarily accumulates in tumors and organs with a high glucose uptake, like the brain or liver.

3.1.1 Radiation physics

Annihilation radiation happens subsequent to the position emission from the nucleus. A positron is an antimatter equivalent of an electron. It is a positively charged particle with the same mass as an electron. It emits from the nucleus of an atom due to radioactive decay. A positron is only emitted from the nucleus as a secondary action to a nuclear reaction, in which a neutrino is created. A neutrino, ν, results from the conversion of a proton into a neutron. The general equation of positron decay is described as

$$ {}^{A}_{Z}X \quad \rightarrow \quad {}^{A}_{Z-1}Y + {}^{0}_{1}\beta^{+} + \nu + Q\,, \tag{3.1} $$

where Q is energy. A proton-rich mother atom, X, loses a proton to achieve stability (daughter atom, Y,). The positive charge of the proton is carried away from the nucleus with a positron, β^{+} [125, p. 342–343]. In Fig. 3.2 the annihilation radiation is sketched. In this case ${}^{18}_{9}\mathrm{F}$ is the proton-rich mother nucleus, which emits a positron and becomes a stable ${}^{18}_{8}\mathrm{O}$ isotope. The positron annihilates outside the atom and two photons are emitted in opposite directions with energies of 511 keV.

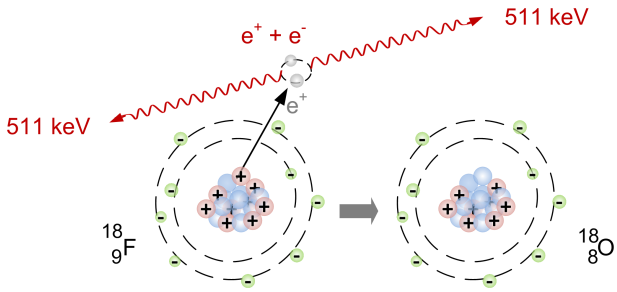


Figure 3.2: β^{+} decay of an ${}^{18}\mathrm{F}$ nucleus with following annihilation.

Certain factors limit the resolution of a PET scanner: one is annihilation outside the atoms, another is uncertainty of the radiation angle of the photons due to scattering, as elaborated in the next subsection. The consequence is reduced image sensitivity and increased radioactive dose for the patients. The path of a positron in matter is random because of elastic and inelastic collisions with anatomic electrons and nuclei. A high-energy positron has a greater possibility

of many collisions, before it finally comes to rest. Clearly this means that it also has a greater possibility of traveling longer, e.g. the mean range in water of a common PET isotope ^{18}F is 0.6 mm. [11, p. 21–23].

Scatter of photons and attenuation
Unfortunately, only 35% of the annihilation reactions in water are estimated to have zero momentum, causing the two photons to radiate at an angle of exactly 180^o to each other. This means that only a minor part of the annihilation events are true events on a LOR.

The type of light scattering depends on photon energy and absorber material. Compton scattering dominates in tissue with photon energies lower than 2 MeV, which is the case in PET (511 keV). Compton scattering is a photon that is ionizing an atom by interacting with a loosely bound orbital electron. The electron is ejected from the atom and the photon is scattered in an angle that is related to its energy loss.

Whenever a photon interacts with matter, it loses energy and the possibility of a given interaction is described by the cross section, σ [cm^2/atom], of the given reaction. The total atomic cross section is given by the sum of all atomic cross sections for all kinds of reactions. On a macroscopic level, the mass attenuation coefficient, μ [cm^2/g], is used. The intensity, $I(x)$, of a parallel photon beam decreases exponentially with the distance, x, in a homogenous medium. An inhomogeneous medium can be divided into homogenous layers of thickness dx', in which the intensity is reduced by dI and -d$I/I = \mu$dx', thus the intensity is given as

$$I(x) = I(0)e^{-\int_0^x \mu \mathrm{d}x'}, \tag{3.2}$$

where $I(0)$ is the intensity of the incoming photon beam. Attenuation[1] depends on the photon energy, the electron density, the atom number of the absorber, and of course the travel length [125, p. 236-237], [11, p. 24–27].

This means that a large portion of the annihilated photons do not reach the detectors and the signal is lost due to attenuation inside the patient. The mass attenuation coefficient for water and tissue is ∼0.1 cm^2/g and the half-value layer is ∼7 cm. This mean that only 1/4 of a beam sent through a 14 cm wide head will get out on the other side.

Coincidence events
The detected coincidence events can be true, random, or scatter coincidence events. True coincidence events are the ones that represent pairs of annihilation photons that have not been interacting with tissue. Random and scatter coincidence events interfere with the true signal; they do not represent an annihilation

[1] Attenuation is the sum of scatter and absorbtion.

event on a line between two detectors. A random event is two detected photons from two different annihilation interactions, which cannot be distinguished from true coincidences, however the number of random events can be estimated on the basis of the count rate. A scatter coincidence event is detection of two photons of which one or both have undergone interaction with matter, so that the direction has changed and the photons have lost energy. Interfering scatter events are limited by elimination of photons with energies less than a certain energy. Only around 5% of the annihilation radiation from a point source without attenuation is detected, which is a measure of the scanner sensitivity. [11, p. 37, 41–42, 95–96].

The number of true counts has to be high to reproduce the tracer distribution; on the other hand the patient dose should be as low as possible. The number of true counts is normally $> 10^7$ counts for a static brain scan. Here the noise effective count rate (NECR) is important; if the noise level is too high, the information of the image gets lost. NECR is expressed by true, T, random, R, and scatter, S, count rate

$$NECR = \frac{T}{1 + S/T + kR/T}. \tag{3.3}$$

k is either 1 or 2 depending on how the random events are corrected. The count rates or the sensitivity of the system to true, random, and scatter coincidence events are functions of: (1) Activity: number of positron annihilation events per second. (2) Geometry sensitivity: the fraction of annihilation photons pairs that reaches the detectors. (3) The fraction of scatter and random photons that reaches the detectors, and (4) detector efficiency: the probability that a photon that reaches the detector is detected. Thus, the NECR is a function of scanner design and patient characteristics (e.g. scatter distribution and dose).

Scatter and true coincidence events increase linearly with the activity, but random coincidence events increase by the activity squared. In Fig. 3.3 the yellow curve increases twice as fast as the two other curves. The curves bend when the yellow curve reaches about 5 M counts, because the scanner system cannot handle more than 5 M counts, which is also the case for the HRRT.

There is an optimum of NECR, so even if the injected patient dose is ignored, the image quality will not be improved by just increasing the dose. In clinical practice a dose of 200–500 MBq ^{18}F-FDG is used (4–5 MBq/kg) [68].

In summary, the isotope annihilates outside the nucleus, thus the registered event does not represent the exact position of the tracer. The possible length of the travel before annihilation depends on the energy of the isotope. There is an optimum for the injected patient dose, which influences the spatial resolution. If the dose is too low there will not be sufficient data for statistics calculations and

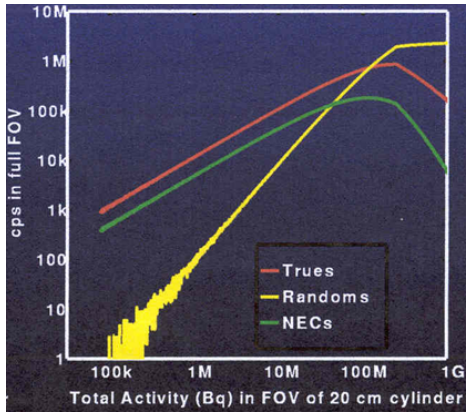


Figure 3.3: Counts per second (cps) as a function of the total activity. NEC is the Noise Effective Counts. On the figure scatter is included in the true curve. At low activities the random events have no effect, so the fraction of scatter is the distance between the red and green curves at low activities [68].

the image will not represent the true tracer distribution. On the other hand, if the dose is too high the noise increases, which also has a negative impact on the resolution. The efficiency of the detectors, how good the detectors are at registering an event, the sensitivity of the geometry, and the path of the FOV that is covered by the detectors, are also important for the scanner performance.

Reconstruction parameters such as matrix size, filter, and use of modeled point-spread-function (PSF) also influence the spatial image resolution.

3.1.2 Scanner physics

Emitted photons are converted to electric signals by a scintillator detection system consisting of scintillators coupled to photomultipliers. The most common and reliable technique couples solid scintillators to photomultiplier tubes (PMT).

A photon deposits its energy in the scintillator that generates a gleam (scintillation), which is transformed to an electric signal and amplified by the PMT.

The scintillator used in PET is a solid crystal where the atoms get excited by particles with energies around 511 keV. Excitation of an atom leaves a hole in the electron configuration, and a free electron fills out this 'hole' in the valence

band, and electromagnetic radiation corresponding to the binding energy is radiated. The number of light photons is approximately proportional to the energy deposited in the crystal, thus the intensity of the scintillator light is a measure of the absorbed photon energy.

3.2 The High Resolution Research Tomograph

The high resolution research tomograph (HRRT) PET scanner is a state-of-the-art dedicated brain PET scanner and therefore this scanner is used for our proof-of-principle experiments. The image quality of the HRRT represents PET scanners of the future. If motion estimation and motion correction can be shown valuable with the HRRT scanner, it will also be valuable for future scanners with their presumably improved resolution.

The design of the HRRT PET scanner is special in three ways; (1) the scintillators consist of two layers of crystals, (2) the number of crystals is 5–10 times higher than for conventional PET[2], and (3) the detector panels are placed in an octagonal geometry (Fig. 3.4(a)). The octagonal geometry is due to the eight detector panels or detector heads. The design of the detector panels is based on a quadrant-sharing design, so the detectors cannot be placed in a ring.

This design is made to improve the detector efficiency and thereby the spatial resolution. One-to-one coupling, where each scintillator crystal is coupled with an individual PMT was used until 1985. The theoretical minimal spatial resolution with one-to-one coupling is equal to the dimension of the PMTs. To improve the spatial resolution, different couplings have been tried. The quadrant-sharing design of the HRRT PET scanner is illustrated in Fig. 3.4(b). Each detector panel is made of 9×13 crystal blocks and 10×14 PMTs. A crystal block has 8×8 crystals in each layer with dimensions 2.1 mm×2.1 mm×10 mm. This gives 119,808 crystals totally in the gantry. The two layers of scintillator material are LSO_{fast} and $LYSO_{slow}$. Between the crystal blocks and the PMTs there is a glass plate to gate the light. The position of the interaction is determined by a weighted sum of the individual PMT signals [30].

The small distance between the detector panels of 35 cm provides high sensitivity. Shortening the length of the LORs minimizes the uncertainty of where the events have occurred, improving the spatial resolution. Similarly, the estimate of the event location is improved with the modern time-of-flight (TOF) reconstruction. The principle in TOF is to limit the event location to a part of the LOR length based on the time of detection between the two affected detectors.

[2]GE Advance PET scanner has 12,096 crystals.

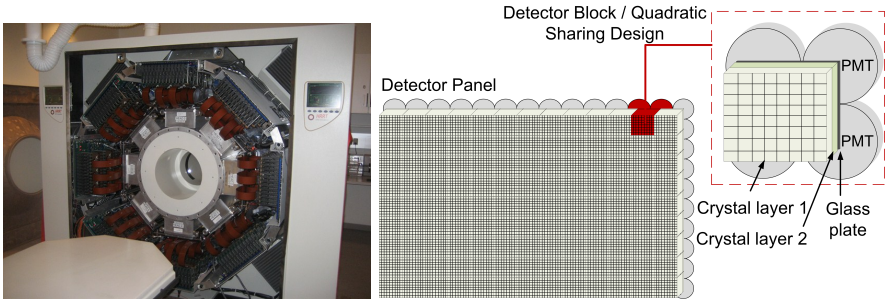


Figure 3.4: Left: The HRRT at Rigshospitalet without front cover. Right: sketch of a detector panel with 9×13 2-layer crystal blocks and 10×14 PMTs. Right: principle of the quadrant sharing design of the detector blocks.

3.3 PET reconstruction

Basically, two kinds of methods for reconstruction exist: analytical by the filtered-back-projection (FBP) and iterative reconstruction (IR). Analytical reconstruction calculates the tracer distribution directly from the measured coincidences. In contrast, iterative reconstruction methods guess at tracer distributions that are iteratively compared to measured data and updated. FBP is a classic and standard algorithm of tomography. It has an easy control of the spatial resolution and noise correlation. This makes it applicable to quantitative image analysis. Iterative methods are less sensitive to the geometry of the detectors (e.g. gaps between the detectors) and non-uniform resolution for the detectors. In this way iterative methods have improved the image quality. [11, p. 63, 71].

There is a variety of IR methods using different cost functions, data models and image models. The two most common methods are the maximum-likelihood expectation maximization (MLEM) and the ordered subset expectation maximization (OSEM). For this project the OSEM 3D reconstruction is used with resolution modeling, which improves the resolution. It uses a simple PSF model that is the same for all the pixels in the FOV and thus it is only a parameter in the OSEM reconstruction [154] and [26].

Figure 3.5 shows a diagram of the OSEM reconstruction with indication by numbers of the corrections prior to the final reconstructed image. The OSEM input is multiple sinograms of rebinned 3D data. Hereunder four correction simograms and the prompt sinogram from the EM list-mode [3]. The sinograms

[3]Every detected event is listed with the time and a pair of detectors represented by coor-

are mapped into a geometric representation of the scanner referred to as the system matrix. The iterative modeling is then applied onto the system matrix represented as the largest red box in Fig. 3.5.

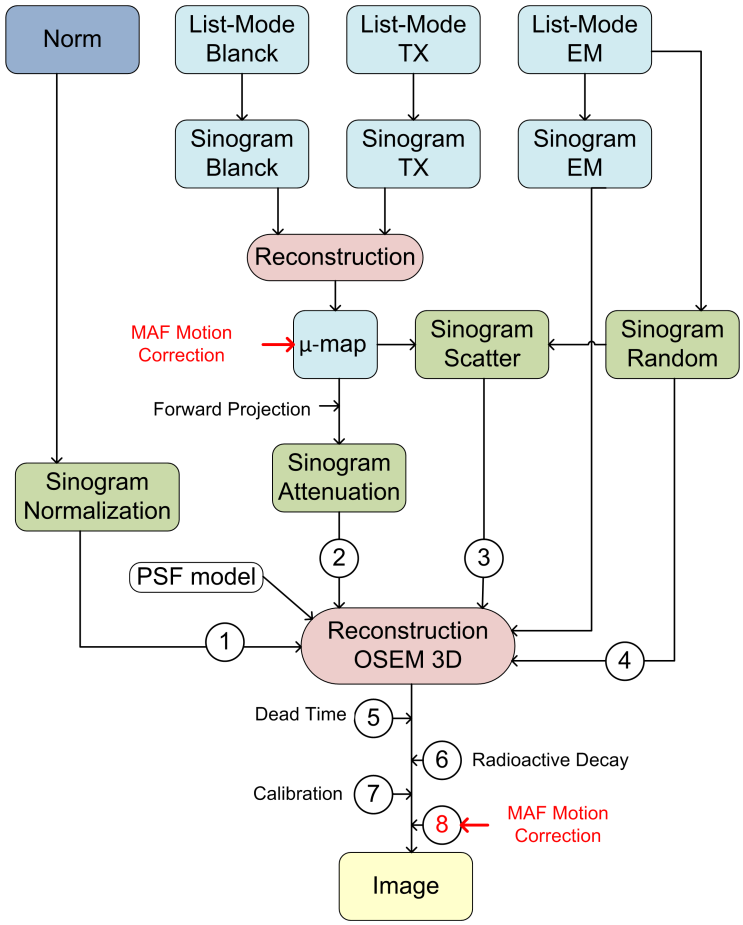


Figure 3.5: Diagram of the 3D OSEM reconstruction used for the HRRT at Rigshospitalet. *Norm* is the normalization. Each number symbolizes a correction operation.

The corrections within the reconstruction scheme are numbered from 1 to 8 and described below including: (1) normalization, (2) attenuation, (3) scatter, (4) randoms, (5) dead time, (6) radioactive decay, (7) calibration and (8) motion correction. The first four correction operations are done within the OSEM reconstruction from the input sinograms. The next three corrections are done

dinates.

on the reconstructed image [30, 85]. The last one in red represents post reconstruction MC with possible repositioning of the μ-map prior to the OSEM reconstruction. The μ-map is a map of attenuation coefficients of the scan subject obtained from the TX scan acquired prior or after the EM acquisition. More advanced MC methods are applied before the IR reconstruction or incorporated within the system matrix of the IR reconstruction model. State-of-the-art MC methods are described in Chapter 4.

Normalization correction - Normalization correction is a compensation of the difference in the efficiency of the LORs. The variation is due to different crystal efficiencies, PMT gains, and positions of the LORs in the FOV. The correction on the HRRT is achieved directly by scans of a homogenous line source to expose each LOR of the same amount of activity [85, p. 74].

Attenuation correction - As mentioned, the photons are attenuated inside the patient. Usually the loss of signal is re-established applying Eq. (3.2) using a μ-map that maps the attenuation coefficient in the FOV. The μ-map is obtained from a TX scan and forward projected into a sinogram [86].

Scatter correction - Scatter events are included in the sinograms, since they cannot be measured. Scatter is minimized anyway by selection of an energy window.

Randoms correction - Random events are minimized using a small time window for coincidence. However, completely avoiding the effect of random events is impossible. The random event rate can be estimated from a measure of the single event rate or by measuring a delayed signal using a delayed coincidence event window. The estimated/measured number of random events for each detector pair is subtracted directly or in a separate sinogram [85].

Dead time correction - Dead time is the time an event takes to be processed, in which new events cannot be detected. Dead time comes both from the dead time of the detector blocks and from the electronics. If photons hit a detector in its dead period the dead time increases. The raw data is corrected for the loss of signal using models calculating the loss of counts of the individual detector pairs expressed by single events [11, p. 116–120].

Radioactive decay correction - During a scan the activity, A_t, of the radionuclide decays exponentially with time, t:

$$A_t = A_0 e^{-\lambda t}\,, \tag{3.4}$$

where A_0 is the activity at $t = 0s$. The loss of signal depends on the decay constant, λ, and the acquisition time.

Dividing a scan into i frames of time intervals, Δt_i the mean tracer activity within a frame, A_i, can be corrected by multiplying a decay factor as

$$A_0 = \frac{\Delta t_i \lambda e^{\lambda t_i}}{1 - e^{-\lambda \Delta t_i}} A_i \,, \tag{3.5}$$

where A_0 is the activity at the time of injection and t_i is the start time of frame i. It is assumed that the tracer distribution is stable within the time interval Δt_i [85].

Calibration - The calibration consists of adjusting the efficiency of the detectors. Sometimes it is included in the normalization term so the normalization reflects the calibration of the detector sensitivities and the geometry of the LOR, i.e. the angle of the incoming photon (LOR-detector angle) and detector-block position (also detector layer for multiple-layer detector scanners such as the HRRT scanner).

3.4 Dynamic PET and 4D reconstruction

In dynamic PET studies, the activity change of the injected radiotracer is measured through sequential time frames. The images of the radiotracer distribution are reconstructed and used with an adequate pharmacokinetic model to estimate the physiological parameters of interest.

The parameters can be tissue perfusion, blood volume, metabolic rates, receptor densities, receptor occupancies, and gene expression rate that have applicable objectives as physiological functionalities or dys-functionalities, or metabolic processes e.g. pharmaceutical behavior.

Most widely used are compartmental modeling techniques applied with TACs extracted from the independently reconstructed image frames. The TACs are extracted from ROIs or from each voxel leading to parametric images.

The parametric images describe the distribution of the physiological parameters and they are obtained indirectly from the PET images in a two-step procedure for the standard 3D imaging. First, the detected events are binned into corresponding sinograms of time frames and reconstructed independently. Second, voxel-based TACs of the time-framed PET images are estimated and applied to a pharmacokinetic model [21].

Advanced 4D PET reconstructions are used for research purposes and increasingly for clinical imaging. They receive increasing attention in the literature,

not the least in relation to the motion correction discussion. The 4D reconstruction methods are differentiated by not reconstructing the time frames independently. An example of one class of method is the direct parametric images reconstruction, in which the pharmacokinetic model is incorporated in the iterative reconstruction process. The first methods were developed by Snyder [149] and Lange and Carson [90] in the 1980s. Recent surveys of direct parametric image reconstructions for PET imaging are given in [161] and [133].

Basically, 4D reconstructions use information from multiple time frames in order to optimize the voxel-by-voxel kinetic analysis. The potential advantages of 4D reconstruction are reduced computation time, improved statistical analysis increasing the signal-to-noise ratio, improved temporal resolution, and thereby the capability to render the dynamic leading to better description of the physiological parameters [21, 133].

As the name suggests, dynamic imaging deals with time-varying intensities reflecting the radiotracer distribution over time. Motion during data acquisition has the same influence on moving image intensities as a function of time. Thus it is crucial to be able to distinguish between these two aspects. The optimal solution to the subject's motion in PET brain imaging is the use of an external tracking system that returns the head pose and the time of the detected events independently of the PET images.

Chapter 4

Motion Correction

Motion correction (MC) of the PET data requires information about the motion to correct for. So a complete motion strategy involves motion tracking (MT) and motion correction (MC) as sketched in Fig. 4.1. This chapter focuses on MC of the PET data (orange), while the next chapter considers the MT part (green). State-of-the-art MC methods are described in this chapter, followed by elaboration of the methods adopted and applied in this project, including the multiple acquisition frames (MAF) MC and the LOR-based motion correction MOLAR reconstruction. The MAF and MOLAR methods are applied in paper A, paper D, and paper E to demonstrate the use of external tracking for PET brain imaging as well as the integration of the tracking systems with the HRRT PET scanner.

An obvious classification of the MC methods is based on the domain or level within the reconstruction the MC is applied. However, many MC strategies performed during the reconstruction can be associated with more domains and are difficult to label. For example if the LORs are corrected to coordinates in the LOR domain whereas the normalization is modeled in the image domain. Figure 3.5 from the reconstruction Section 3.3 demonstrates the complexity of all the input to the iterative OESM process with supplementary calibration and dead time correction afterwards. Thus the motion correction is often applied in several domains. In the following we distinguish between MC within the reconstruction, referred to as LOR/sinogram domain approaches and MC of the images after the reconstruction, referred to as image-domain approaches (Fig. 4.1).

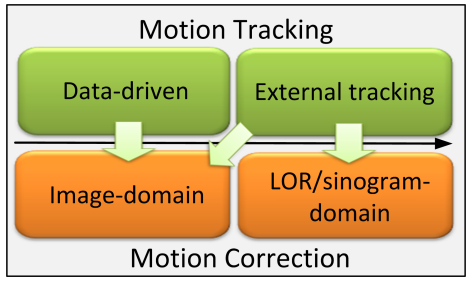


Figure 4.1: Motion tracking (MT) can be data-driven or estimated with an external tracking system (green boxes). This motion information is used as input for motion correction (MC) of the PET data, either post reconstruction or within the reconstruction referred to as image-domain or LOR/Sinogram-domain respectively (orange boxes). External MT can be used for all kinds of motion corrections, while data-driven MT has insufficient time-resolution for LOR-based MC.

4.1 State-of-the-art: motion correction in PET brain imaging

A wide range of MC algorithms have been proposed over the last two decades for PET imaging, of which brain imaging constitutes the largest share [14,16,22,45, 89,100,123,127,130,137,167]. This is explained by an increased interest in brain imaging and perhaps more likely the simplification of the motion problem with respect to the brain. Hence the MC methods are often demonstrated on the brain and then adopted for non-linear cardiac and respiratory motion [89,137]. It is accepted to consider the brain and skull as one immobilized body, static in shape and structure, so head motion represents rigid body transformation of the brain. Thus intracranial movements during the scan are ignored.

Image-domain MC occurs post reconstruction. One method is the image co-registration, where the EM frames are reconstructed independently and re-positioned to a reference position. This method was proposed by Picard and Thompson [123] referred to as the multiple acquisition frames (MAF) method and is described in more detail in the next section. The MAF method has since been evaluated on phantoms and human studies and has shown to improve the PET studies compared to no MC [35,45,109,121]. The MAF method is widely used nowadays because of its simplicity, despite its limitations [35,109]. The MAF method assumes that the normalization is unchanged and suffers from lack of correction of intra-frame motion. In addition, if iterative reconstruction is applied, short time frames can be problematic since they can converge toward a bias [17].

An alternative post reconstruction MC method involves deconvolution techniques of the image frames with mathematical operators expressing the measured motion [39, 100]. The approach is to reduce the motion-induced image blur. However, the deconvolution process may enhance the noise of the PET images as described in [132].

LOR-based MC corrects the individual LORs into the position as if no motion had occurred [14,16,22,29,44,100]. List-mode correction has a theoretical potential of accurately correcting for all kinds of motion. The cited LOR MC studies above concluded on two weaknesses of the LOR reposition method, which resulted in artifacts on the corrected images. The two weaknesses were described in [158] and [157]. First, the scatter distribution changes after the realignment, and second some of the data cannot be used, because some events are now out of FOV after realignment. These missing events reduce the signal and thereby the NECR and also introduce artifact if the missing events are not compensated for.

These problems are more significant on the HRRT scanner with an orthogonal scanner geometry resulting in gaps and thereby increasing the ratio of events exiting the FOV or events entering the FOV due to motion. Rahmim and colleagues also developed a method to incorporate these fictive motion-induced events in the reconstruction of the HRRT scanner [132].

In addition, especially in recent years, MC of the normalization or sensitivity factor has caused much discussion in the literature due to complexity from motion [10, 130, 138].

A reasonable amount of counts is needed for proper normalization estimation of all individual detector efficiencies. In theory, the normalization should account for all possible directions and times. This is especially a problem for scanners with many detectors, like the HRRT, and with expected number of LORs less than one on average for all the possible $4.5 \cdot 10^9$ detector pairs.

An early approach to overcome the normalization problem was to use the normalization of the repositioned detector pairs and not the detectors actually affected [100], introducing image artifacts [14]. Later on different approaches have been suggested to model or simplify the normalization [10, 130, 138].

In general LOR MC is used for pseudo-LOR MC in the sense that correction of all individual LORs is a computation-heavy task, basically impractical few years ago [22]. Often the LOR are rebinned into subsets of small time duration [102, 130], and [69] or spatial subsets where neighboring LOR are grouped [14] and corrected within histogram space or the system matrix. Hong and Reader showed that histogram-based MC is much more time-consuming

than list-mode-based MC [69].

An LOR rebinning approach with time-bins defined by a motion threshold has been applied for correction within the system matrix [130] and for 4D reconstruction with the use of a time-varying normalization approach [162]. The use of a motion threshold is similar to the ones suggested for the MAF method [45,123], [109]. A motion threshold is not a well-defined measure, since the influence on a specific ROI varies with the distance from the origo and should be used with caution.

Nowadays, with still increasing computational power and improving scanner technologies, 4D reconstruction with incorporation of MC for all LORs is having increasing impact and seems to be the dynamic reconstruction of the future [133, 135].

4.1.1 Multiple acquisition frames

The MAF method is often used for MC comparison studies [78,102] and proof-of-principle studies where MC is only a sub-method of a larger MC strategy, as is the case for paper A, paper D, and also in [83,109]. In paper A and paper D the MAF method is used to demonstrate integrated PET acquisition and external tracking recorded with the Polaris Vicra and the Tracoline system.

The MAF method uses multiple frames, partitioned according to movement magnitude, such that the intraframe movement is below some threshold. Frames are reconstructed independently and aligned according to dynamic registered motion during the PET scan [123] (Fig. 4.2).

List-mode data are divided into frames after a given threshold of movement. The TX scan cannot be divided into frames as the radiation source moves as a function of time, thus this method assumes that there are no movements during the TX scan. This is not only the case for the MAF method, but for all correction schemes with the HRRT PET scanner. Before the individual EM time frames (A_n) are reconstructed, the μ-map ($A_{\mu n}$) is rotated to the position of the head at the given EM frame (Fig. 4.2). There are two ways of aligning the μ-map, either the TX image can be rotated and used in the reconstruction, or the detectors can be "moved" before the TX scan is reconstructed. In this way the scanner is conceptually moved instead of the patient. When the EM images are reconstructed, using the different aligned μ-maps, the EM images are rotated into a chosen reference frame [45, 64, 121]. In [109] we proposed to carefully select the reference position in order to reduce the number of reconstructions, including problematic short frame reconstructions. If a central

position is chosen as the reference the number of detected motions might be reduced, resulting in fewer frame reconstructions and reducing the computation time and interpolation errors.

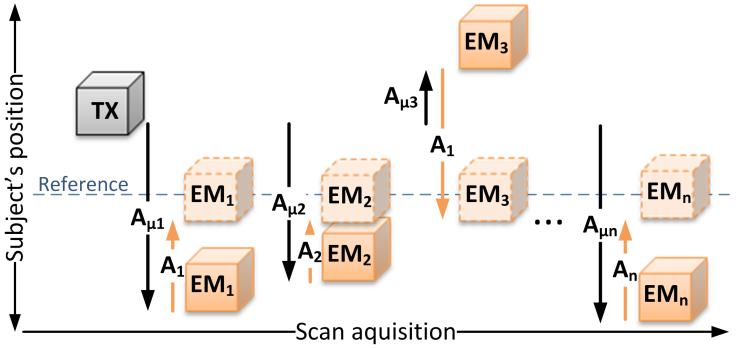


Figure 4.2: Schematic drawing of the MAF MC procedure. Repositioning of the μ-maps ($A_{\mu i}$) before reconstruction and realignment of the individual frames (A_i) to a reference position.

Assuming the exact 3D head position is known, the reconstructed images are repositioned by 3D rigid transformation as elaborated in paper A. Mesh grids of the original image coordinates are transformed into the reference position. The original images are then interpolated from the rotated coordinates into the original coordinates to avoid empty pixels. In paper A and paper D tri-linear interpolation is applied.

4.1.2 MOLAR

The motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction named MOLAR is specially designed for the HRRT PET scanner [22]. This MC method was used as the basis in paper E to evaluate the performance of the Tracoline tracking system for rigid MC in PET brain imaging.

MOLAR is a list-mode event-by-event MC algorithm developed to achieve the highest possible spatial scanner resolution permitted by the computation time permits. A few approximations to perfect solutions were made to make the algorithm practically usable. The most significant approximation is related to the global sensitivity term including the normalization.

The MOLAR reconstruction is designed to work with the Polaris Vicra tracking

system and assumes the head motion is perfectly registered and transformed into the scanner coordinate system. The LORs are rebinned into time-bins corresponding to the sampling rate of the Polaris system (50 ms).

The MC is performed on the LORs following modeling within the system matrix possible for iterative reconstructions. The system matrix presents the geometry of the scanner and expresses each voxel as the radial and axial voxel-to-LOR distance. LORs are mapped into the system matrix on the basis of the motion input from the Polaris system. The geometry of the LOR is defined by motion input, while the associated voxel contribution along the corrected LOR is modeled by the iterative OSEM algorithm, as briefly described in Section 3.3. This modeling occurs iteratively within the system matrix, including the expected normalization, random, and scatter factors [22]. The attenuation is fixed from a backward mapping of the data to the TX scan position. Normalization and randoms estimates are related to the actual detectors involved in each coincidence.

The LOR sampling differs from frame to frame, hence the normalization is varying over time. This makes computation of the global sensitivity term intractable. The global sensitivity term is related to the attenuation and the normalization of the PET events with respect to space and time [22, 130].

In the MOLAR algorithm the normalization is modeled based on a selection of detector LORs accounting for motion [10, 138]. An unique algorithm normalization term is calculated for each frame based on the motion data.

Rahmim et al. [130, 132] simplify the global sensitivity term by pre-correction for the attenuation and model the motion in the system-matrix domain.

4.2 Motion correction strategies at HRRT scanner sites

There are seventeen HRRT PET installations worldwide. The majority are used for research purposes exclusively (Fig. 4.3). At Rigshospitalet, Copenhagen the HRRT scanner is used for research subjects and patients.

The MOLAR MC algorithm was the first MC method developed for the HRRT and also the only MC strategy with external MT implemented for many years. In recent years, several of the seventeen sites with an HRRT installation have published MC-related work on the HRRT due to the fact that the MC is needed to utilize the high resolution of the scanner [22,27,35,39,78,83,109,130,145,162].

Figure 4.3: There are eighteen HRRT PET scanners worldwide, of which seventeen are considered as a custom HRRT sites. The remaining scanner is located at Siemens, Knoxville.

MOLAR has been implemented at the National Institute of Health (NIH) and the Yale PET center by Carson and colleagues for research purposes and it is routinely used. The University of British Columbia, Vancouver and McGill University, Montreal have also now demonstrated MC methods beyond MC in the image-domain [130, 162].

The HRRT User Community has provided MC software for the HRRT sites, including the MAF method integrated with selected motion input, either 1) data-driven automated image registration (AIR), or 2) External MT motion input with the Polaris Vicra system as example. The software provided was partly based on the presented implementation of the Polaris system with the HRRT PET scanner at Rigshospitalet in paper A. The software was tested in [83] and [109].

Most implemented MC strategies at the HRRT sites use the Polaris Vicra system combined with data-driven MT (AIR implemented in user software or customized versions). It has repeatedly been reported at users' meetings that this marker-based system potentially causes incorrectly registered head motion due to the uncertainty of the marker attachment to the head. It is difficult and basically impractical to validate if the MT with the Polaris system is correct using the data itself or the image data.

Danielle and colleagues have suggested making a quality control (QC) of the registered motion by comparison of motion registered using AIR and Polaris [35]. This QC approach can limit the use of incorrect motion input for the MC but it does not solve the problem.

Chapter 5

Motion Tracking

This chapter concerns the first part of a motion compensation strategy in PET brain imaging. How can motions be registered? Previously suggested motion tracking (MT) methods are briefly described in this chapter, including a more detailed description of the popular Polaris Vicra system.

5.1 State-of-the-art: motion tracking in PET brain imaging

Accurate head motion registration during the PET acquisition is essential for correct MC in PET brain imaging. Basically, two classes of methods have been proposed; 1) motion estimated from the image data itself, referred to as a data-driven approach, and 2) external MT obtained from other technologies using alternative hardware devices (Fig. 4.1 in Chapter 4).

In data-driven methods the list-mode EM acquisition is rebinned into time frames, as known from dynamic PET studies. A pose is obtained from each time frame, representing a head motion during the time frame. In general, pre-defined framing based on the dynamic study protocols is used. Several methods exist to obtain the motion between reconstructed image frames based on image co-registration [74, 77, 97, 120, 148, 153, 168]. These methods are widely used for co-registration of multi-image modalities such as PET to magnetic reso-

nance imaging (MRI) image registration. Mutual information (MI) [97,153] and automated image registration (AIR) are popular [169] methods in PET brain imaging.

Data-driven MT prior to image reconstruction has been suggested [41, 59, 145]. These techniques have not really had any impact due to the substantial amount of data statistics required, which are not present in PET with a low EM event rate. In order to utilize MC methods beyond image-domain approaches, motion capture of higher time-resolution is required; a time-resolution that is better than what is possible with image-driven MT methods.

This means that, along with the MC method developed for PET imaging, external MT methods have also been developed. Basically, the proposed methods register an object that is assumed to be fixed to the patient's head. An early approach was stereoscopical detectors that register a rigid body with three light emitting diodes (LEDs) attached to the patient's head [55]. Along with the MAF method, a stereo camera system capturing motion of three LEDs attached directly to the patient's head was developed assuming a rigid constellation [122]. A stereo camera system registering a self-encoded patterns was developed for for MRI [42]. A single camera system for narrow scanner geometries viewing a 3D solid tool with markers was presented in [103]. The mentioned systems were customized and only demonstrated on few studies.

Similar commercial systems are also available and commonly used in functional MRI (fMRI), SPECT, and PET brain imaging [12, 93, 100, 170]. These systems archive stereographic perspectives from multiple cameras of an optical target attached to the patient. A commercial stereo camera system registering a checkerboard pattern has been demonstrated on awake animal PET studies [88]. The preferred systems in PET brain imaging are Polaris systems from Northern Digital Ins. (NDI) [93], of which the Vicra model with a small view angle is generally used for narrow scanner geometries [22, 39, 116, 130].

Markerless MT systems have been recommended several times and mentioned as an obvious solution to avoid time-consuming and failure-based attachments of marker objects to the patient's head or animals [6, 38, 72, 87, 96, 106]. None of these demonstrate motion correction in PET imaging. It is non-trivial to achieve accurate and robust markerless tracking and this remains to be developed and demonstrated in medical brain imaging.

A revolutionary methodology with simultaneous MRI and PET acquisitions has recently been presented for human scanning [31]. One advantage of this image multimodality is the opportunity to incorporate MRI sequences for motion estimation within the PET reconstruction [23, 160].

5.1.1 Advantages of external MT over data-driven MT

The suggested MC methods for PET brain imaging assume precise motion measurement of the head. In the previous chapter we distinguished between LOR/sinogram MC and image-domain MC (Fig. 4.1). For the latter, data-driven MT methods can be applied assuming insignificant variation of the radiotracer distribution. This is of course extremely critical in dynamic PET imaging, since the time-varying tracer distribution is exactly the aspect of interest. A major advantage of external MT is that the motion is estimated independently of the EM data.

A second important advantage of external MT over data-driven MT is the much higher time resolution. LOR-based MC reconstructions require information of the head pose with a time resolution adequate to capture motion of approximately 50–200 ms. Accurate pose estimation during such small time samples is impossible to achieve from the EM events themselves due to insufficient statistical basis stemming from low count rates. Thus an external MT system is required for LOR-based MC.

5.1.2 The Polaris Vicra system

The Polaris Vicra is an optical real-time tracking system. The Polaris system includes two main components; a sensor and a set of rigid tools. The position sensor registers 1–6 tools at a time and returns the position(s) and orientation(s) in quaternions. In Fig. 5.1 the position sensor is placed behind the HRRT facing the patient tunnel. There are two kinds of tools - passive or active. The active tool consists of infrared (IR) light sources, unlike the passive one. The passive one only reflects IR light and has a simple setup compared to the active, preferable in PET imaging. A passive tool consists of 3–6 spherical retro-reflective markers. The maximum tracking frequency in passive mode is 20 Hz, which is more than sufficient to capture human motion during PET acquisition. The position sensor provides an IR source for passive markers, collects their data, calculates their pose and sends the results to a host computer. The position sensor consists of two charge coupled devices (CCDs), each surrounded by an illuminator ring of infrared light emitting diodes with centre wavelengths of 880 nm. When infrared light reaches the markers, some of the light is reflected back into the position sensors. The CCDs detect the intensity and pattern of the light, and by pattern recognition and triangulation techniques the Polaris Vicra system returns the 3D position and orientation of the tool. The tracking is done simultaneously with a 3D root mean square (RMS) accuracy of the translation of 0.25 mm [4].

Figure 5.1: The Polaris position sensor placed behind the HRRT's patient tunnel and the scanner bed is in the front of the HRRT.

As default the position is measured relative to the center of the position sensor, although it is possible to define one of the tools as the reference point. There are some advantages in using a reference tool fixed on the HRRT. It makes the system more flexible and stable, because the position sensor does not have to be placed at the same spot for each measurement, nor does it even have to be at exactly the same position during a scan, (although this is preferable). Therefore the position sensor can be removed between scans without any extra calibration.

5.1.3 Discussion

The major disadvantage with the Polaris system for PET brain imaging is how to firmly fix the tracking tool to a patient's head? It will always be a source of error and precise head motion estimation is crucial for MC in PET brain imaging. Many different approaches to overcome the problem have been tried. Examples are shown in Fig. 5.2. The one that is most widely used and perhaps also the one with most success is a cap with additional fixation using straps and bandages as demonstrated on the image to the right. This approach has been used at the Yale PET Center with success in many research studies. However, it is an impractical solution for clinical use. Further in research studies it is often impossible to prove that the attachment did not fail, which could be significant for the study conclusion.

Figure 5.2: Different tool attachments. A) Goggles used in Turku PET Center, Finland. B) Cap applied in Manchester, UK (however they have disapproved the approach and test goggles now). C) Our suggested hygienic method used at Rigshospitalet (but with many flaws). D) The method applied at the Yale PET Center, New Haven, USA.

A second problem with the Polaris system for PET imaging is the time synchronization. The system does not provide a gated signal for time-synchronization with other hardware devices. The Polaris does not return a time related to the tracking frames, it only reports the poses related to a frame number. The file creation time of the tracked poses is often used as time reference.

Chapter 6

Alternative Tracking Methodologies

The next two chapters are an introduction to the developed Tracoline system. This chapter discusses alternative head tracking methodologies for PET brain imaging.

6.1 Requirements

A head motion tracking system needs to satisfy multiple technical and clinical requirements, if it is to be applied with PET scanners in clinical settings. The clinical requirements are regardless of the MC method applied, in contrast to the technical ones. In order to be used for hospitals the tracking needs to be:

1. Time-saving
2. Easy to use
3. Comfortable for the patients
4. Hygienic
5. Reliable

Diagnosis and treatments have to be based on valid foundations. Thus a reliable MT is crucial, irrespective of the MC method used. A tracking system must have an easy interface with the PET scanner and it must also be sturdy and

stable to be part of the daily routine in the hospital. Further, it must not be uncomfortable, since an uncomfortable patient will introduce movements which, even though correctable on the images, are counter-productive for both patient's well-being and image quality. Finally, the hygienic conditions of hospital use have to be fulfilled. It must not be a source of infection. Items in contact with patients must be washable or disposable.

The technical requirements are:

1. Sub-millimeter precision
2. Suitable sample frequency
3. Suitable measuring volume
4. Small viewing angle
5. No interference with the PET acquisition
6. 3D real-time registration (preferable)

The accuracy of the tracking system has to be better than the spatial resolution of the PET scanner, otherwise the MC will increase the blur instead of reducing it. The sample frequency has to be at least twice as high as the head motion frequency in order to avoid aliasing, according to the Nyquist criterion. However, this is only a consideration for continuing motion during the scanning. A few high-frequency motions are generally not important due to the low number of EM events in a short time frame [93].

The field-of-view of the tracking system has to cover the expected measuring volume of the patients. The geometry of the HRRT scanner is especially challenging for integration of an external MT, since the patient tunnel is narrow and limits the viewing angle of a potential registration sensor. Figure 6.1 shows a side-view drawing of the HRRT geometry, the suggested position for the tracking device illustrates the limited field-of-view. Optimally, the registration of the positions should be estimated simultaneously and stored with the list-mode data, so the MC reconstruction can be processed immediately. The workflow is highly important for modern medical scanners.

6.2 Alternative tracking methods

Multiple tracking solutions exist based on different methodologies, e.g. magnetic [66, 128], optical [46, 93], radio-frequency [9], global positioning system (GPS) [37], accelerometers and gyroscopes [71, 95], and camera-vision-system [20]. Several surveys of computer vision systems have been published, a few elaborate ones are given in [47, 101, 126].

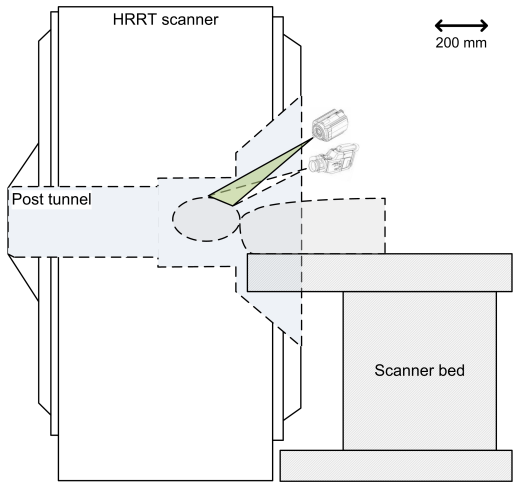


Figure 6.1: Drawing of the HRRT scanner to scale. Two tracking devices are pointing at the patient's face region demonstrating the limited view.

In order to satisfy the requirements in Section 6.1, a tracking system should not include any markers attached to the patient. Hence the technologies listed above are limited to vision systems, including time-of-flight (TOF) cameras. These systems have the advantage of being without object attachments.

The TOF camera technology is an optical-camera-sensoring technology. From a single camera, 3D information is obtained of a scene with a high frame rate necessary for tracking. Amplitude-modulated light is sent towards a scene and the reflected patterns are captured by a CCD sensor. The phase-shift of the emitted light and the incoming light reflects the depth of the scene [91]. The absolute accuracy of the depth measure is typically less than one centimeter [107] [1]. Thus this technology is problematic for pose tracking in PET brain imaging, due to the required tracking accuracy of sub-millimeter.

Multiple camera views is a second way to obtain 3D information without using markers. The problem with pure camera systems is to find robust and accurate correspondence between the camera perspectives. In general, feature points are identified in the different images followed by a procedure matching the extracted features. Scale-invariant features (SIFT) extraction is a popular method for automatic feature extraction. To each feature a SIFT descriptor is assigned like a "fingerprint". The descriptors are used to solve the matching problem between extracted feature points in different image planes [94]. To solve the ambigui-

[1] A leading company of TOF cameras, MESA Imaging AG, Zürich, Switzerland claim an absolute accuracy of ±10 mm for the SwissRanger^{TM} SR4000 camera.

ties, camera-based tracking systems often use passive or active markers that can be identified on the video stream. Such video-based systems are commercially available and widely used for motion tracking within medical imaging and radiotherapy (e.g. RPMTM system, *Varian Medical Systems Inc.*, Palo Alto, USA, MicronTracker, *Claron Technology Inc.*, Toronto, CA, and systems from *Vicon*, Oxford, UK,) [88, 119]. Optical systems such as the Polaris Vicra similarly use CCD sensors to capture light reflectance of markers.

Unique matching of feature points without the use of markers can be obtained by replacing one camera with a light source. The light source represents a view perspective as the camera but the light can be recognized in the other image-plane(s), resulting in more robust 3D measurements [80].

Multiple surface scanning systems exist with varying types of light sources, cameras, and principles of reconstruction [24]. Over the last decades , some of the methods have been commercialized (i.e. systems from: *GFMesstechnik GmbH*, Teltow, Germany, *3shape*, Copenhagen, DK, *Canfield*, Fairfield, USA, *3dMD*, Guildford, UK, *L-1 identity solution*, Billerica, USA, and *visionRT*, London, UK). These are examples of companies providing surface scanners for face reconstruction. The methodology has rapidly improved over the last few years, reaching an acquisition speed which now makes surface tracking possible. Recently, visionRT presented systems for gated CT and radiotherapy treatment. Microsoft Corporation has also recently launched the popular Kinect gaming platform. This is an imaging system demonstrated for human tracking in [146]. The system uses structured light to create a depth map explained in Section 6.3.

Surface scanner systems are highly application-specific and existing systems do not fulfill all the requirements for PET brain tracking with respect to: geometry, speed, optics, accuracy, and robustness.

Surface scanners using structured light have improved significantly with the development of digital light processing (DLP) [43]. This technology was developed in 1987 and has since improved displays and projectors considerably. The technology is based on digital micro-mirror devices (DMD) consisting of an array of micro-mirrors with one or a few mirrors representing each pixel of the display. The mirrors are switchable between two discrete angular positions with illumination 'on' and 'off'. Graduation of pixel intensities is achieved by integrating over a time period [15]. The first hand-held Pico projector was commercialized in 2008, just before the beginning of this project. The miniaturized design and the improved DLP image quality regarding sharpness and contrast made surface scanning possible in PET brain imaging. However, the Pico projectors available use visible light incompatible with head tacking during PET scans. It would be uncomfortable for the patients to have light projected onto their face during long PET acquisition.

During this project the first Pico projector was re-engineered to use invisible (short-wave infrared) light [112].

6.3 Structured light 3D surface imaging

3D imaging covers techniques that reflect the true psychical world, e.g. the well-known medical imaging modalities PET, MRI, SPECT, and CT. These techniques reconstruct the physical body in volumetric image arrays. Alternatively, 3D surface imaging represents 3D point coordinates on a surface in 3D space. A depth map is another 3D representation, also used in surface imaging. A depth map is described as a function of a two-dimensional space, e.g. from a view point or a plane.

Structured light 3D surface imaging is a class of methods where the surface is reconstructed using a light source and a sensor. The principle is demonstrated in Fig. 6.2. An active illumination source modulates a pattern of spatial variation onto a scene. An image sensor projects the irradiated scene onto a 2D image plane (e.g. camera, CCD chip). Information on the 3D surface is obtained from the distortion of the illuminated pattern, assuming a known model between the illumination source and sensor, referred to as system calibration. In Fig. 6.2 a point on the target is denoted Q and the corresponding points in the two perspective image planes are given as q_c and q_p for a camera and a projector respectively. The 3D coordinates for Q can be obtained from simple triangulation.

Several kinds of illumination patterns have been suggested to achieve correspondence between the illumination plane and the sensor image plane in order to get 3D point coordinates of the irradiated surface. One method is the projection of a sequence of patterns that binary codes the illumination plane that is recognized in the image plane with black and white stripes. This method is highly robust but requires many patterns. Instead of only binary codes, a number of gray-scale values can be utilized; referred to as a Gray code [8] [141]. Furthermore, colors can be incorporated in the coding to increase the number of code levels. In Fig. 6.2, a rainbow pattern is used for illustration. Each color is uniquely related to a line position of the illumination plane [49]. Fast methods to incorporate more patterns into one-projection utilizing the RGB colors were proposed by Huang [73] and Geng [50, 51]. Different structures such as stripes, grids, dots, pseudo-random segments, and combinations are explored in the literature. A comprehensive survey of existing methods is found in [48].

A structured light 3D surface system is highly application-specific, depending on

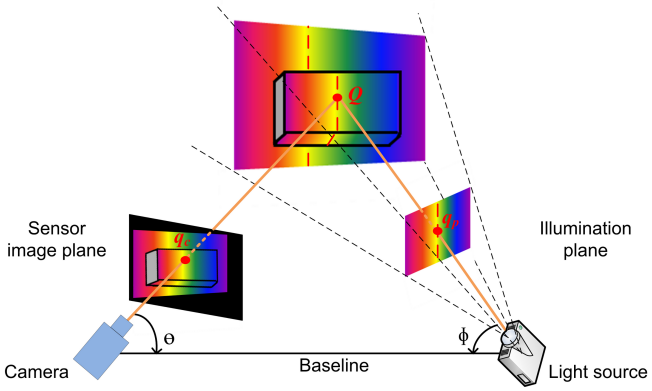


Figure 6.2: The principle of structured light systems. A light source is illuminating a rainbow pattern onto an object (box) and a camera is capturing the scene. Note how the projected pattern is distorted when it interferes with the object.

the required accuracy, spatial resolution, speed, robustness, and optical scene. To be used for motion tracking, a suitable speed is required and therefore a minimum of patterns is preferred. Fast methods with only one pattern to obtain a 3D surface are possible from most of the coding principals listed above. However, this is at the expense of accuracy and robustness.

For face tracking in medical imaging the optical scene is challenging due to the surface properties of the skin, the limited view, and the restriction of the illumination which cannot be visible. The use of wavelengths above the visible region reduces the sensitivity of CCD cameras, resulting in reduced image contrast. Basically, methods relying on discrimination of the pixel intensities are inappropriate. This is also supported by the properties of the skin, which reduce the contrast since light scattering is prevalent on a semi-transparent surface such as the skin [159]. In addition the skin is inhomogeneous due to moles, hair, scars, etc. Hence the method has to be robust against color and contrast variations.

6.4 Phase-shift interferometry

Phase shift interferometry (PSI) is a technique from optical metrology adopted for structured light 3D surface imaging to archive correspondence between image perspectives using phase-shifted cosine-patterns [19, 142]. The advantages of this method are that; (1) it is robust toward intensity variation between the

fringes and along the fringes. (2) Theoretically, it is possible to gain sub-pixel resolution due to the fundamentally analogue cosine function. (3) It is a fast method, considering it as a robust method; only three patterns are needed to obtain a surface.

Phase-shift interferometry and fringes projection are frequently mentioned in relation to the use of cosine-patterns and also stripe-patterns, for structured light imaging.

In conventional interferometry, fringe projection is obtained from optical interference between light beams resulting in a static interference pattern referred to as interference fringe.

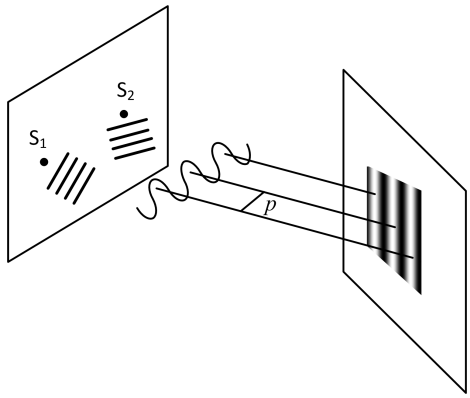


Figure 6.3: Principal of phase-shift interferometry. Two coherent light sources (S_1, S_2) with same intensity and same center phase but shifted creating an interference fringe that is seen on a projection surface. This illustration shows particularly the cosine interference fringe from two planar wavefronts. P is the fringe spacing corresponding to the spatial frequency.

The intensity distribution of an interference fringe is visible on a surface whenever stationarity of the two beams occurs. This means that the relative phase between the two wavefronts is constant. A stable interference fringe of a cosine pattern is achieved from two coherent light sources producing planar wavefronts with a constant phase-shift as illustrated in Fig. 6.3 [19] [2]. The intensity distribution across a surface from an interference fringe $I(x, y)$ can be captured by a camera referred to as an interferogram (Fig. 6.3). In PSI the intensity of the

[2]Incoherent light sources such as white light can generate interference fringes if a static field is achieved. However, with less expression of the wavefront, resulting in fringe pattern of bright and dark regions, which is why, stripe-patterns is also referred to as fringes projection. In principle, phase and amplitude modeling of the light beams can result in a wide range of fringe patterns.

interferogram is measured while varying the phase difference between the two incident waves creating exact information of the surface. In Fig. 6.4 an interferogram is captured by a camera. The axis of depth is defined by the camera and the angle between the fringe projection and the camera is α. $f(x, y)$ is the hight of the surface. P is the fringe spacing and is given as a function of the angle and wavelength of the two incident beams of equal intensity I_0. The spatial phase modulation function $\psi(x, y)$ is then:

$$\psi(x, y) = \frac{f(x, y) \sin \alpha}{P} \tag{6.1}$$

The interferogram from the camera view can be expressed by the spatial phase modulation function as

$$I(x, y) = 2I_0 \left[1 + \cos\left(\frac{2\pi x}{P} + \psi(x, y)\right)\right] \tag{6.2}$$

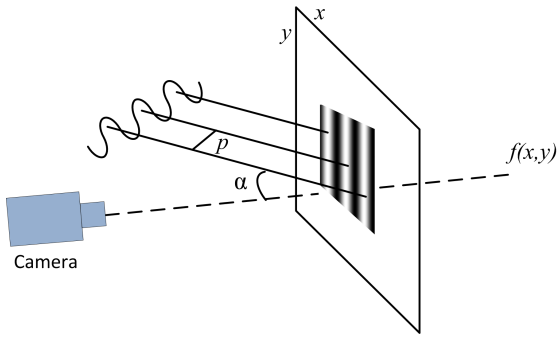


Figure 6.4: A camera is capturing an interferogram from a fringe projection source.

A shape change of the surface is described by the change of the spatial phase modulation function equals a phase $\varphi(x, y)$ ($\varphi(x, y) = \psi_1(x, y) - \psi_2(x, y)$). Hence a relative phase-shift between an object and a reference beam is equivalent to the height of the object [144, 150, 156].

A light projector can simulate a given interference fringe such as two planar waves by simple projection of a cosine-pattern. Thus a light projector can substitute advanced optical laser systems in conjunction with an interferometer[3] for surface imaging [73]. The advantages of light projectors over optical interferometers are the flexibility of pattern modulation and control of the projection

[3] An interferometer generates interference fringes projection by, for example, splitting the incident beam and generating a phase difference via differentiating the propagation lengths of the two beams.

rate. Optical interferometers can produce highly coherent interference fringes with analogue interferograms, while light projectors digitalize the pattern. However, the DLP technology has improved the light projection considerably for this purpose.

The three-stage method [28] with projection of phase-shifted cosine-patterns is widely used to achieve depth information and correspondence between the projector and camera images [43, 73, 171]. The captured interferogram in the camera space (u_c, v_c) can be derived from Eq.(6.2) and is given as a function of three unknown parameters; the average intensity background $I_{\text{av}}(u_c, v_c)$, the spatial intensity modulation $I_{\text{mod}}(u_c, v_c)$, and the phase $\varphi(u_c, v_c)$:

$$I(u_c, v_c) = I_{\text{av}}(u_c, v_c) + I_{\text{mod}}(u_c, v_c) \cos\left[\varphi(u_c, v_c) + s\right] \tag{6.3}$$

where s is a phase-shift. The phase that describes the depth of the object can be found from projection of three uniformly phase-shifted cosine-patterns $I(u_p, v_p)$, i.e.

$$I(u_p, v_p) = a\left[1 + \cos\left(\frac{2\pi}{P}u_p + s\right)\right] + b, \quad \text{and} \quad s = \begin{bmatrix} \frac{2\pi}{3} & 0 & -\frac{2\pi}{3} \end{bmatrix} \tag{6.4}$$

where (u_p, v_p) refers to the projection image plane, a is the amplitude, and b is the bias of the cosine function. The phase can then be obtained from three phase-shifted cosine interferograms solving Eq. (6.3) with respect to the phase:

$$\varphi(u_c, v_c) = \tan^{-1}\left(\sqrt{3}\frac{I_1(u_c, v_c) - I_3(u_c, v_c)}{2I_2(u_c, v_c) - I_1(u_c, v_c) - I_3(u_c, v_c)}\right) \tag{6.5}$$

The phase directly gives a depth-map of the projected surface, which can be converted into a distance measure from a calibration object with known geometry. However a phase-unwrap is required to achieve unique correspondence between phases and heights due to a 2π angular period of radians. The unwrapping procedure is illustrated in Fig. 6.5 for the undistorted projected pattern (left) and the distorted pattern captured in the camera (right). The periods along the first dimension of the angular phase are added to a multiple of 2π archiving a unique depth-map $\varphi'(u_c, v_c)$.

The depth-map can be converted to 3D point coordinates from correspondence between the projection image plane and the camera image plane given a calibrated system. The image plane correspondence is in principle the depth-map from the distorted and un-distorted fringe patterns demonstrated by profiles in Fig. 6.5 last row.

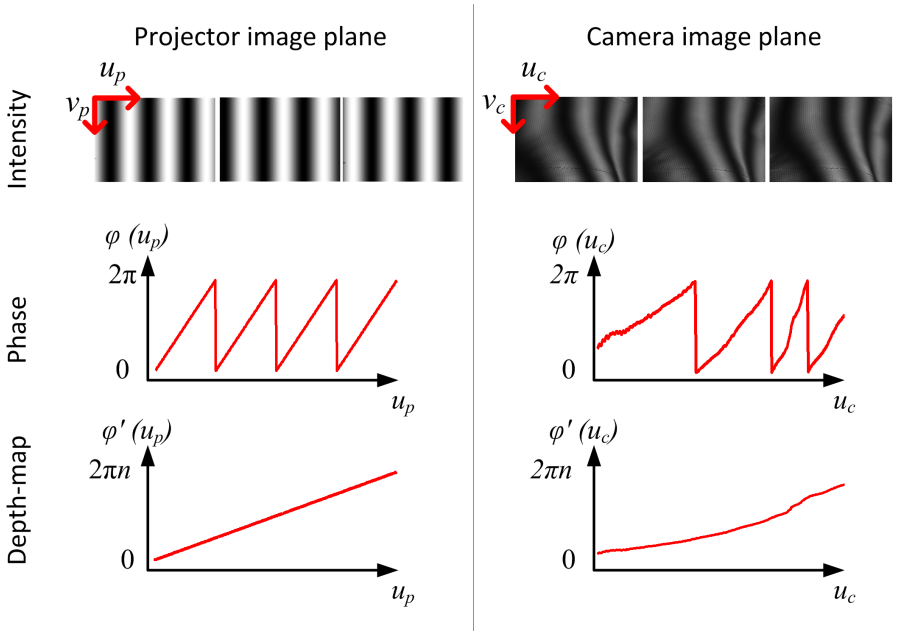


Figure 6.5: Schematic procedure of the three-stage method to obtain correspondence between a projection plane and a camera image plane.

Chapter 7

"Tracoline" - A Surface Tracking System

The Tracoline system is an outcome of this PhD project. The system is a surface scanner that is capable of capturing projected patterns with a frame rate of 30 fps with two cameras. This is utilized to generate 3D point clouds of a moving surface and thereby track the motion of the surface. This chapter describes the principles of the system not elaborated in the papers presented in the appendices.

7.1 Hardware components and control

The main component of the Tracoline system is a DLP Pico Development Kit from Texas Instrumentation. This product is a compact DLP Pico projector made for custom-engineered systems and products. The Kit was delivered with interface to enterprise volume management system (EVMS) that recognizes a small number of systems excluding windows PCs (version 2.0 is capable of this).

The DLP Pico development kit was chosen due to its flexibility. It has a synchronization signal output which is used for camera synchronization and the design also made it possible to change the light sources. GFMesstechnik GmbH also started incorporating the DLP Pico Development Kit into their product line and made a GFM Pico Developer Kit available that includes a small Linux

EVMS control board and application programmers interface (API) for window PC and source code to control the DLP Pico projector. The GFM control board was built into the Tracoline system.

Tracoline is based on DLP technology from Texas Instruments and uses phase-shift interferometry (PSI) to archive correspondence between camera and projector similar to the GFMesstechnik surface scanning products (i.g. PRIMOS, FaceSCAN 3D, and MicroCAD) [43].

However, Tracoline is special with the use of near infrared (NIR) light and is the first and only surface scanner with Pico DLP technology using non-visible light. The original LEDs of the Pico projector have been replaced with a near infrared emitting diode (NIRED). The characteristics of the Pico projector with the in-built NIRED are described in the technical report in Appendix F. For this prototype the NIRED is cooled by a fan mounted above the Pico projector [1].

Figure 7.1 shows a diagram of the Tracoline system in the gray box with the DLP Pico projector centrally located.

Timing: Two FLEA2 cameras from Point Grey Research (PGR) are triggered via a synchronization signal output from the control board at the beginning of a new projected pattern. The signal is directed from the DLP Pico projector. The captured images are time-stamped using embedded image time stamps based on the firewire clock running at 8000 Hz (an IEEE 1394 cycle timer) [75]. It is the recommended time-stamp mechanism for PGR cameras. Timing is important for a tracking system, especially for synchronization with other hardware devices such as a PET scanner.

System control: The PGR cameras are firewire connected to a host PC used for data storage and running the system control program. The system control program is developed in C++ and is based on PGR and GFM APIs to run the cameras and Pico projector, respectively. The system control program was made by Michael K. Rasmussen in a sub-project to this PhD project [134]. An important part of running the Tracoline system in tracking mode is error detection. A projected pattern also has to be captured, stored, and used for pose estimation. For example if the images cannot be written to disk fast enough, the images within the camera buffer start getting overwritten. The error detection occurs at two steps; 1) the internal cameras' counters are checked for discontinuities. 2) Time between a newly captured image and the previous image is constantly checked against the projection frame rate.

The communication within the system control program is indicated with black

[1] the in-build of the NIRED was done by an external company.

lines in Fig. 7.1, whereas the hardware components and connections are in red. The blue part represents the possessing of the raw image data referred to as the calculator. This could be built-in to the Tracoline system. However, during the development process it has been post-processed as shown on the camera line to the left in Fig. 7.1.

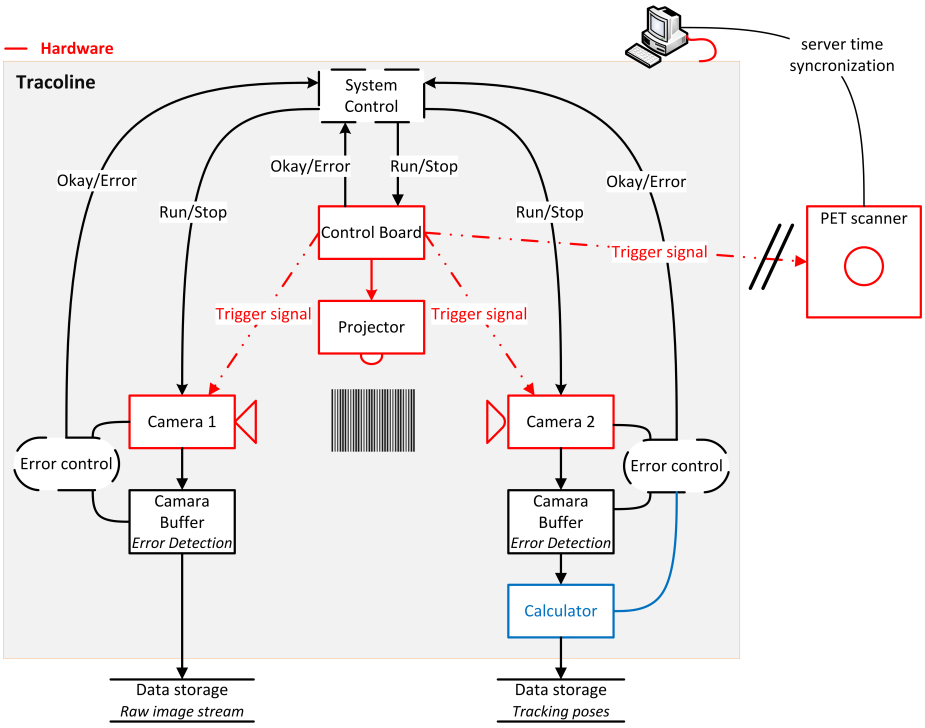


Figure 7.1: Diagram of the Tracoline system (inside the gray box) and its internal and external communication paths. Red components represent hardware. The blue part represents the processing of the pose estimation from the raw images.

7.2 Positioning and processing of raw image data

The system is preliminarily designed for the HRRT scanner and the only possible position for an SL system is in the front as seen in Fig. 6.1 due to the long post-patient tunnel. We have chosen the upper face region as target, especially

around the bridge of the noise. In this region curvatures are dominant and facial movements are limited, rendering it optimal for surface alignment. Two cameras are preferred to accommodate the challenging viewing angle inside the scanner tunnel which introduces occlusions, mainly from the patient's nose.

The tracking output is based on surface-to-surface alignments of point clouds of the face region. The point clouds are obtained from PSI by fringes projection of cosine-patterns using the three stage method described in Section 6.4. The cameras and projector are internally calibrated using the perspective pinhole model to achieve two sets of point clouds in 3D coordinates.

In paper B we present an initial procedure, considering the projector as a pseudo-camera, warping captured images onto the projector image plane [172]. Then using the popular method from [173] with projections from a set of 2D planes often used for stereo camera calibration. Later, the calibration procedure was optimized to determine the intrinsic and extrinsic calibration parameters from the model in [173] directly without inaccurate image warping. The optimized procedure is based on projection of a grid of feature points onto a plane at different views expanding the FOV captured by the cameras.

The further pose estimation from the 3D face surfaces is elaborated in paper C.

7.3 Integration of external motion tracking with PET motion correction

External MT in PET brain imaging has to be integrated with the PET scanner in order to be applied for motion correction [45, 123]. An MT system has to be installed physically at the scanner without interfering with other equipment and personnel around the scanner. The scanner rooms are well stocked with equipment and there are many activities around the scanner during operation. A small mobile MT device is a must. Further, the system has to be synchronized with the PET scanner with respect to time and geometry.

A complete PET brain motion correction scheme with incorporation of an external MT is seen in Fig. 7.2. The scheme is threefold; initialization, data acquisition, and data processing. The red boxes are a special case using the Tracoline system to visualize the geometrical calibration procedure and the pose estimation principle. The initialization involves synchronizations with the HRRT scanner and is elaborated below. The blue part of Fig. 7.2 represents the simultaneous head tracking and PET scanning. The collected data are post processed

(green part). The PET image reconstruction is created incorporating the motion information demonstrated with repositioning of LORs.

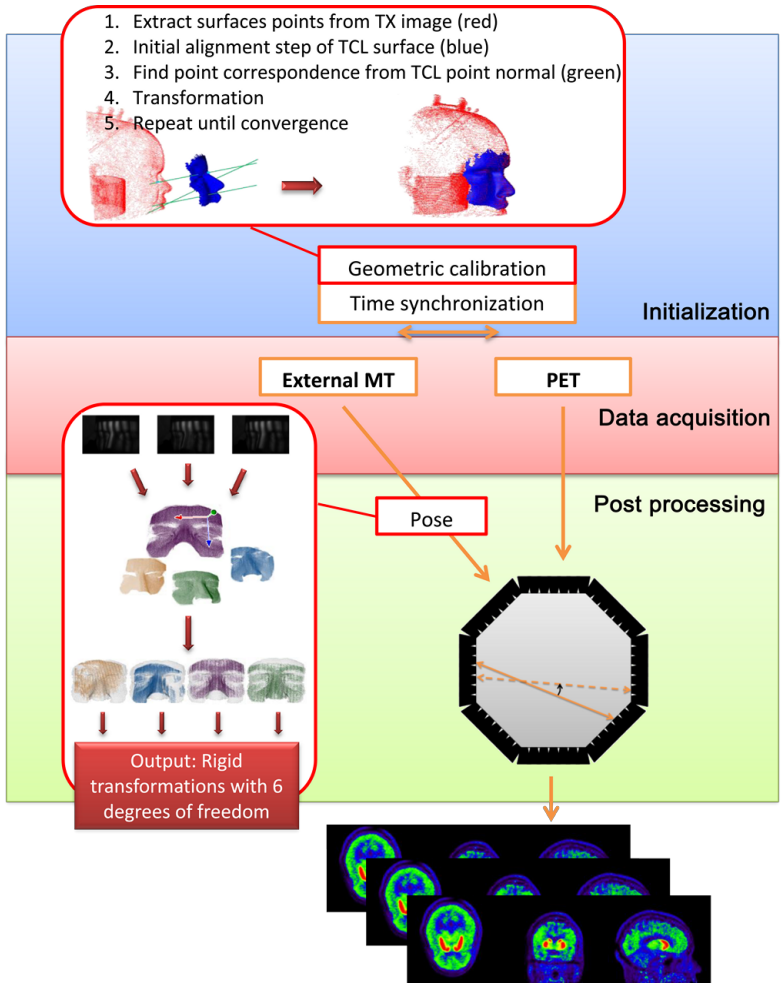


Figure 7.2: Motion correction scheme with incorporation of external tracking information. Blue part: initialization involves time and geometric synchronizations of the MT system and the PET scanner. Red part: Head motion recorded during the PET acquisition. Green part: PET reconstruction incorporating the motion information. The result is motion-corrected PET image-frames seen at the bottom. The highlighted red boxes represent operations related to the Tracoline system. (1) Upper box shows how the geometric calibration/synchronization is obtained, (2) Lower box shows the pipeline of the head pose estimation acquired from the raw-video stream.

Geometric calibration: The papers presented D and E involve integration of the Tracoline system at the HRRT scanner at Rigshospitalet and the Yale PET center, respectively. In the proposed calibration method, the affine transformation matrix is found from correspondence between the TX image volume and the reference surface of the Tracoline system. In Fig. 7.2 the scan subject is a mannequin head. However the same principle was used in paper E on human scans. The method is described further in paper D and E. An alternative method was demonstrated in [84]. In this approach surfaces from the TX image volume and the Tracoline system are aligned by an iterative closest point (ICP) algorithm [166].

Time synchronization: The Tracoline system returns the embedded timestamp from the firewire clock relative to the first tracking frame after initialization of the system. At this first tracking frame, the local time of the host PC is also returned. Thus synchronizing other hardware devised with the host PC is a proper solution. It should be noted that it is not recommended to use the absolute time of the host PC, since it is often unstable and less accurate [79] [2]. It is possible and more precise to synchronize external devices with the Tracoline system via the synchronization signal output from the control board, as visualized in Fig. 7.2. The first solution was used for this project. In PET scans with minor high frequency motion, time-synchronization within a second is adequate in conjunction with the relative low rate of the EM data collection. However it is not preferable, since sudden motion is often seen in human brain PET. Accurate time-synchronization is not an issue for MC methods post-reconstruction.

Incorporation of tracking information: The PET image frames or LORs have to be repositioned in the same coordinate system used for the external tracking. Alternatively, the tracking information must be transformed into the HRRT coordinate system. In principal, it is the same; a resulting final transformation that corrects the PET data at a given time into a reference position. It is a combination of four transformations. In Euclidean space using homogeneous coordinates, a rigid transformation is described by multiplication of a 4×4 matrix consisting of rotation and translation parts. A sequential transformation is a matrix multiplication. Thus the complete transformation $\mathbf{A}_{\text{total}}$ sequence is a multiplication of four sub transformations:

$$\mathbf{A}_{\text{total}} = \mathbf{A}_{\text{align}}\mathbf{A}_{\text{ref}}\mathbf{A}(t)^{-1}\mathbf{A}_{\text{align}}^{-1}, \tag{7.1}$$

where $\mathbf{A}_{\text{align}}$ is the geometric calibration transformation, and $\mathbf{A}_{\text{ref}}$ and $\mathbf{A}(t)$ are output transformations from the tracking system at the reference position and at a given time, respectively. Irrespective of whether an image-based or an

[2] Synchronization of PCs against a timeserver (e.g. network time protocol (NTP)) might for example not be stable since clock drift occurs depending on timeserver, lag time on the internet and constrains of the operation system.

LOR-based motion correction algorithm is applied, the correction to a reference position is a sequence of transformations, as further described in paper A and paper E. The rotation can also be applied using quaternion representation. In the papers presented in the appendices, the Euler angles are used due to simple matrix operations in Eq. (7.1).

7.4 Safety certification

The Tracoline system is a customized development with in-built optics. The combined near infrared emitting diode (NIRED) and the Pico projector is a customized combination.

A risk assessment and classification is to be done with the final product and not the individual components. The built-in components are characterized by the producers. However, the characteristics change when integrated.

A safety certification has to be completed before the system can be applied for human face surface reconstruction. The Tracoline system has to comply with the laser standard (IEC-60825) [3] from the International Electrotechnical Commission (IEC). The maximum permissible exposure (MPE) and accessible emission limit (AEL) for a class 1 laser are:

$$MPE = 639 \pm 81\mathrm{W/m}^2 \quad \text{and} \quad AEL = 25 \pm 3\mathrm{mW}$$

The system has been characterized through a series of experiments as described in a technical report found in Appendix F. A third party has reproduced the results presented in the technical rapport and certificated the system. The certification document is in Appendix G. The documented maximum exposure and emission for the Tracoline system are 3.02 $\pm$44 W/m^2 and 6.43 mW, respectively. These values are much lower than the limits for a class 1 laser. Thus the Tracoline system can be applied for face measurements without exposing the eyes or skin to any hazard.

7.5 PET studies

A substantial contribution of this project is a method and a prototype for collection of PET data with simultaneous external MT(s). Prior to data collection the external MT systems have been integrated with the PET scanners by mounting

the system on the PET scanners, a geometric calibration procedure, and timing synchronization protocols.

Data presented in paper A and paper D of phantoms were acquired on the HRRT PET scanner at Rigshospitalet. Additionally, six HRRT PET scans of patients with simultaneous Tracoline tracking have been obtained at Rigshospitalet during the development period.

Data on a human volunteer and phantom presented in paper E were acquired during a stay at the Yale PET center with Professor R. E. Carson and his group. At the Yale PET center, advanced list-mode MC was available. List-mode MC is a more challenging mode to test a tracking system in PET brain imaging, in contrast to the MAF method which is the only method available at Rigshospitalet.

An additional four HRRT PET scans of volunteers with simultaneous Tracoline and Polaris tracking have been obtained at the Yale PET center. However, these studies were not reconstructed with the Tracoline motion input, since they were incorrect due to interference from the Polaris system.

Chapter 8

Conclusion

The main novelty of this project was the demonstration of a markerless tracking system with an accuracy improving the quality of PET brain images. During the project a markerless tracking system was designed, constructed, tested, and integrated with a PET brain scanner.

Head motion is a significant problem in functional brain imaging. Many commercial head tracking systems exist and are widely used for research purposes. However, they rely on markers attached to the patient's head, resulting in potentially unreliable head tracking.

Markerless head tracking is of high interest and several research groups have suggested using markerless tracking. Markerless tracking is nontrivial for PET brain application due to the requirements of precise and robust pose estimation. Thus markerless head tracking has not previously been demonstrated with PET acquisition.

The design was made for the HRRT PET scanner representing scanners with narrow geometries. An accurate structured light 3D surface imaging methodology was preferred for the application. The design of the system was based on state-of-the-art hardware components, further improved for the application. A miniaturized DLP projector was modified to use near infrared light, and the cameras and optics were adjusted for the challenging field-of-view. The design was based on two cameras to limit the risk of occlusion causing lack of pose estimation. The Tracoline system can handle occlusion of one of the cameras and still capture the motion.

To accommodate the need for fast and accurate tracking, digitalized phase-shift interferometry was chosen for generation of 3D point clouds and surfaces. Pose estimation was based on state-of-the-art surface-to-surface alignment. The root-mean-square (RMS) distance of the resulting alignment was in the sub-millimeter range for both phantom and human surface scans (mean of the RMS point distances between the aligned 3D point cloud and the reference surface).

The accuracy of the Tracoline system was tested against a rotary stage and mechanical measured translations. The motions measured were compared to the results of a commercial tracking system popular in PET brain imaging. The two tracking systems had similar accuracy tracking a rigid object. Rotary and translatory accuracies were measured to RMS errors of 0.09 degrees for ± 20 degrees axial rotations, and RMS errors of 0.26 mm for ± 25 mm axial translations.

The optics of the Tracoline system were characterized and documented safe for face tracking in described settings.

The Tracoline system was integrated into two HRRT scanner sites (Rigshospitalet and the Yale PET center) and significant improvement of motion correction was shown on phantom studies. PET motion correction based on the Tracoline was shown to be similar to motion correction based on an integrated tracking system, both on phantom and human studies.

Tomographic scanners with integrated motion tracking systems do not exist. Yet, such a scanner system has the potential of directly associating a motion descriptor to a registered PET line-of-response. Furthermore, building a tracking system into the tomographic scanner optimizes two important aspects: precision and automatization. Time synchronization and geometric calibration can be improved and automated.

Appendix A

A Moveable Phantom Design for Quantitative Evaluation of Motion Correction Studies on High Resolution PET Scanners

Oline V. Olesen, Claus Svarer, Merence Sibomana, Sune H. Keller, Søren Holm, Jørgen A. Jensen, Flemming Andersen, and Liselotte Højgaard

Abstract

Head movements during brain imaging using high resolution positron emission tomography (PET) impair the image quality which, along with the improvement of the spatial resolution of PET scanners in general, raises the importance of motion correction. Here we present a new design for an automatic, moveable, mechanical PET phantom to simulate patients' head movements while being scanned. This can be used for evaluating motion correction methods. A low-cost phantom controlled by a rotary stage motor was built and tested for axial rotations of 1°–10° with the multiple acquisition frame (MAF) method. The phantom is able to perform stepwise and continuous axial rotations with sub-millimeter accuracy, and the movements are repeatable. The scans were acquired on the High Resolution Research Tomograph (HRRT) dedicated brain

scanner. The scans were reconstructed with the new 3D ordered subset expectation maximization algorithm with modeling of the point spread function (3DOSEM-PSF), and they were corrected for motions based on external tracking information using the Polaris Vicra realtime stereo motion-tracking system. The new automatic, moveable phantom has a robust design and is a potential quality assessment tool for development and evaluation of future motion correction methods.

A.1 Introduction

Positron emission tomography (PET) imaging has changed the impact and standards of nuclear medicine. In the last decade PET image quality has improved considerably and comprehensive use of radiochemistry with new advanced PET technology has further developed the methodology. Patient motion lowers image quality, especially for high resolution PET scanners. The High Resolution Research Tomograph (HRRT, Siemens) is a brain dedicated scanner, with a resolution down to 1.4 mm when using a new 3D ordered subset expectation maximization (OSEM) reconstruction algorithm with resolution modeling [115]. This method incorporates a spatially invariant point spread function (PSF) [26,154].

The three main kinds of head motion relevant for PET scans are: long drift motion, frequent short motions around a mean position, and occasional quick movements. The most common kind of head motion is long drift motion, occurring when patients are relaxing. Significant motions are observed when patients move their heads up and down or sideways (leading to rotations around the transaxial or body axes) or when they move their legs (leading to axial translation of the head) [34,140]. If there is no correction for patient head motions, they degrade image quality, especially during long acquisitions. They can also cause distortions of functional analysis due to artifacts from inaccurate attenuation and scatter correction or wrongly detected signals in a region of interest (ROI). Head motions of only a few millimeters have a blurring effect on PET brain images taken on a conventional PET scanner with a spatial resolution of 0.5–1.0 cm [58]. This "blurring effect" increases with improving scanner resolution and thus head motions can counteract the technological advances of these high resolution scanners.

To remove the blurring and distortions from head motions, different approaches to motion correction using tracking information have been suggested. These include multiple acquisition frames (MAF) involving repositioning reconstructed frames [45,123], post-reconstruction image based deconvolution of overall motion [39], incorporation of the overall measured motion within the system matrix

of the reconstruction algorithm [131,137], and motion correction of individually detected events [16,100,167], with the addition of motion based normalization correction [14] or modification of the system matrix [22,132]. Combinations of methods where sinograms are corrected within each frame, after which the frames are repositioned have also been reported [102]. These correction methods are in general evaluated using a regular phantom such as a multi-line-source, hot-spot or 3D Hoffman phantom or a mathematical phantom, where movements are computer simulated and/or performed manually.

In this paper we present our design, implementation, and evaluation of a newly developed phantom with coiled tubes filled with radiotracer. The phantom can simulate the most common movements of a patient's head during scans and more importantly, the simulated movements are repeatable. A repeatable mechanical phantom setup makes it possible not only to compare different correction methods, but also to evaluate the source of errors of any correction method and the benefits of a correction method for a given scan protocol or type of motion.

Our hypothesis is that the optimal choice of motion correction method depends on the type of motion and this hypothesis has motivated the design of the phantom. The phantom can create ground truth motion of a highly complex, brain-like structure and it can enable future evaluation of when and how to correct for patient motions in PET brain imaging.

The aim of this study was therefore to develop and test a new mechanical, moveable PET phantom as a tool for future improvement of motion correction in PET scanners.

A.2 Materials and Methods

A.2.1 PET

All scans were performed on an HRRT dedicated brain PET scanner [165] and all images were reconstructed using the new 3D-OSEM PSF reconstruction algorithm [26,154], improving the resolution of the HRRT scanner down to 1.4 mm [115].

A.2.2 Motion Tracking

An optical 3D motion tracking system, Polaris Vicra (Northern Digital Inc.), was used to register movements during the scans. The system uses a rigid tracking tool with 3–6 markers, and the system measures the orientation and position of the tool using a quaternion representation. A position registered by the Polaris must correspond to a known position in the HRRT in order to reposition a given line of response (LOR) or frame correctly. This means that the transformation between the coordinate system of the HRRT image frame and the coordinate system of the Polaris tracker coordinate system has to be established (Fig. A.1). Alignment between the tracker coordinate system and the image frame only needs to be estimated once, if the Polaris sensor is fixed and does not move over time relative to the gantry. If the Polaris is not fixed, a reference tool fixed to the gantry has to be used. We have used a customized reference tool placed inside the back end of the tunnel close to the patient tool (Fig. A.2(a)). The Polaris output was set to give the position and orientation of the patient tool relative to the reference tool, assuming the reference tool was fixed to the HRRT gantry [45, 129]. This means that the origin of the tracker coordinate system was the position of the reference tool in contrast to what is shown in Fig. A.1, where the origin is at the sensor (default setting). The advantages of using a reference tool are that you can move the Polaris sensor and have free access to the scanner subject from the back of the scanner, e.g. when tubes or devices are put through the back end of the tunnel. Moreover, you can move the sensor between scans, and in theory during a scan, although this is not recommended. Moving the sensor before a new scan can be necessary as the sensor does not always 'see' the patient tool in every position.

Using a reference tool, the measured alignment refers to the 3D rigid transformation that transforms the reference tool into the image frame (transformation between the tracker coordinate system and the image frame).

In [45] and [129] the alignment is obtained using paired emission (EM) scans and Polaris trackings of a point source or a line source placed relative to a tracking tool. We preferred to use transmission (TX) scans in order to avoid handling radioactive material. We used paired high statistics transmission (TX) scans and Polaris trackings of the patient tool to have direct point correspondence between the tracker coordinate system and the image frame [14].

The alignment transformation was determined by eight such paired measurements of TX scans and Polaris trackings of the patient tool with four markers. The 3D rigid transformation between the two coordinate systems was found by identifying the same set of points (the markers of the tool) and then using the closed-form loop solution to estimate the absolute transformation [129]. An au-

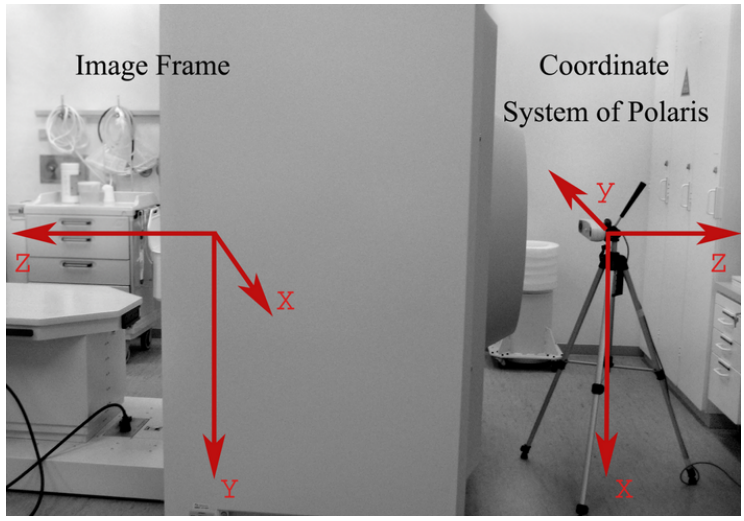


Figure A.1: Polaris Vicra sensor on a tripod behind the HRRT PET scanner. The two coordinate systems are indicated. Note that the image frame by convention is left-handed (we have reversed its y-direction to get two right handed coordinate systems).

tomatic routine developed for this purpose was used to identify the markers on the TX images. The algorithm used the center of mass assuming that the top of the screw was located in the center of the spherical marker and that the tool was placed in a transaxial image plane. Prior knowledge of the markers' positions relative to each other was used to improve the estimated marker position. The TX scans of the patient tool were obtained with high statistics to improve the image quality.

A.2.3 Phantom Design

A moveable phantom was designed to simulate head movements during PET scans. Figure A.3 shows a diagram of the phantom; a photograph of the phantom assembly is shown in Fig. A.4. The body of the phantom is an acrylic plastic cylinder (D = 180 mm, L = 165 mm) in which two 152 cm long connector tubes (LPDCT160, Medrad Performance) were coiled up and fixed by glue to create a complex 3D structure (volume 5 ml). The inner diameter of the tubes was only 1.5 mm to create small details for the high resolution scanner. The connector tubes represent a brain and could be filled with e.g. an ^{18}F-fluorodeoxyglucose (FDG) solution. The connector tubes were rigid to avoid bending that could

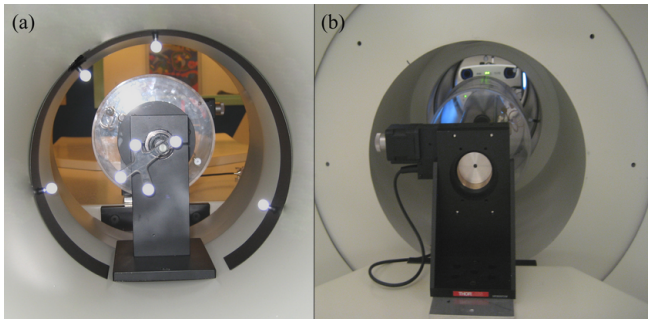


Figure A.2: Back (a) and front (b) view of the phantom setup inside the HRRT scanner. The markers used by the Polaris Vicra tracking system for both the reference tool and the patient tool placed by the phantom are clearly seen as the bright white spots in panel (a).

block off the tracer flow.

The axis of rotation was positioned in the center of the cylinder. Three carbon pipes (D = 6 mm, L = 370 mm) were fixed on a metal mounting plate at each end. The carbon pipes were placed with a triangular geometry with a distance of 70 mm between the pipes to stabilize the axis of rotation. One of the mounting plates was fixed to a rotary motor (NR360S/M, Thorlabs) connected to a controller (BSC102, Thorlabs) and a computer. The stage controller software (ATPUser Utility software) programmed the stage to rotate stepwise or continuously with a given velocity and acceleration. The opposite mounting plate was connected to a combi-ball bearing (RAKB-16, Sanisteel) by a 142 mm long axis at the center in the opposite direction of the carbon pipes. The phantom could rotate around the axis and move back and forward without friction within the combi-ball bearing, as indicated in Fig. A.4. Thus, using the motor stage the phantom could make well controlled movements to simulate sideways rotation of a patient's head around the axial axis, and by moving the scanner bed horizontally the phantom could translate along the axial axis, the type of motion seen in human PET scans when subjects move their legs. Stepwise or continuous slow movements could be simulated with a resolution of 1 arcsec and a max velocity of $15°/s$.

Eight-mm thick perspex plates with holes for the carbon pipes and with inlets and outlets for the radioactive solution were placed at the top and bottom of the cylinder. A screw mounting was fitted at the end of the axis of rotation to fasten the Polaris tracking tool. In Fig. A.2 the complete phantom setup is shown inside the HRRT scanner. The two tubes could be filled to simulate different activity in grey and white matter in the brain. The end of the phantom

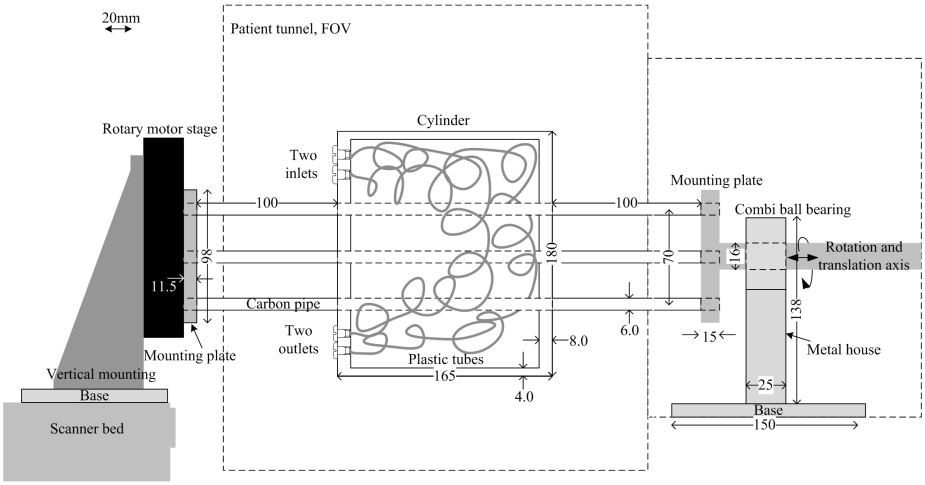


Figure A.3: Side view drawing of the moveable phantom in place inside a small field of view PET scanner.

with the ball bearing was placed inside the back of the patient tunnel and the end with the motor was placed on the scanner bed. Only the cylinder was located inside the field-of-view (FOV) and this part of the phantom was made of materials with low attenuation to eliminate the need for attenuation and scatter correction. Although the phantom design is simple, it has a detailed 3D structure. All equipment is available on the market and the phantom was built in collaboration with the Mechanical Workshop at Rigshospitalet.

A.2.4 Phantom Experiments

One 11-minute stationary EM scan of the phantom was obtained and reconstructed in frames of 60 s using the 3D-OSEM-PSF (16 subsets and 10 iterations) algorithm assuming no attenuation. Eight EM scans of the phantom with an acquisition time of 15 minutes each were performed while the stage was moving the phantom and the Polaris system was registering the motion. The rotary stage was programmed to rotate the phantom by 10 degrees in steps of 1 degree with a dwell time of 60 s at each stationary position. The stage was used to achieve identical motion for all 8 scans. The tubes of the phantom were filled with an ^{18}F-FDG solution. The activity within the tubes was 10–15 MBq during all scans.

The EM list-mode data were divided into frames based on the eleven positions

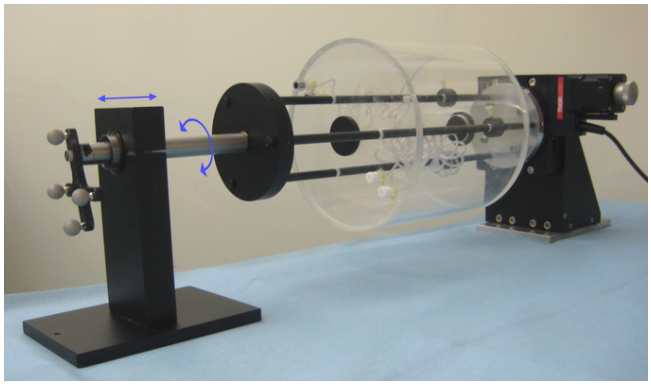


Figure A.4: Side view of the moveable phantom with the tracking tool fixed at the end of the rotation axis. The blue arrows indicate the rotational and translational motions the phantom is able to perform with the combi-ball bearing.

registered by the Polaris system and reconstructed. The frames were repositioned to a reference position using the external tracking information and the previously obtained alignment transformation. This is referred to as the MAF motion correction method. The reference position in each of the eight scans was chosen as the first frame (at 0 degrees) of that specific scan. The images were transformed into the tracker coordinate system, where they were repositioned into the reference position and finally transformed back to the image frame of the HRRT (Fig. A.5). The method is described mathematically as follows.

The PET image is a grid of points which is either in the image frame (x, y, z) or the tracker coordinate system $(\bar{x}, \bar{y}, \bar{z})$ expressed in homogenous coordinates. Thus a fourth dimension is introduced in the vector and matrix representations in order make it possible to use simple matrix calculations for the subsequent transformations in projection space. This is common in 3D transformation mathematics. For simplicity, we consider the patient tool to be a point P. The Polaris system returns the orientation $\mathbf{R}$ and translation $\mathbf{T}$ of the tool in the tracker coordinate system. The translation matrix can be calculated as

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.1}$$

and the Polaris system returns the rotation or orientation of the tool as a unit quaternion q, which can be thought of as an extended complex number with three imaginary parts

$$q = q_w + iq_x + jq_y + kq_z. \tag{A.2}$$

The quaternions can be converted to the Euclidian rotation matrix as

$$\mathbf{R} = \begin{bmatrix} q_w^2 + q_x^2 - q_y^2 - q_z^2 & 2(q_x q_y - q_w q_z) & 2(q_x q_z + q_w q_y) & 0 \\ 2(q_x q_y + q_w q_z) & q_w^2 - q_x^2 + q_y^2 - q_z^2 & 2(q_y q_z - q_w q_x) & 0 \\ 2(q_x q_z - q_w q_y) & 2(q_z q_y + q_w q_x) & q_w^2 - q_x^2 - q_y^2 + q_z^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.3}$$

The common Euclidian rigid transformation matrix can then be calculated as

$$\mathbf{A} = \mathbf{RT}. \tag{A.4}$$

The output of the Polaris is the transformation that moves the reference tool/origin O into the position of the patient tool

$$P'(\bar{x}, \bar{y}, \bar{z}) = \mathbf{A}_{\text{ref}} O(\bar{x}, \bar{y}, \bar{z}) \tag{A.5}$$

$$P(\bar{x}, \bar{y}, \bar{z}) = \mathbf{A}_{\text{out}} O(\bar{x}, \bar{y}, \bar{z}) \tag{A.6}$$

where $\mathbf{A}_{\text{ref}}$ and $\mathbf{A}_{\text{out}}$ are the recorded transformations of the patient tool at the reference position P' (frame 1) and a "new" position P (frames 2–11). Correction of the "new" point into the reference position is performed in the tracker coordinate system by combining (A.5) and (A.5) (as illustrated within the big box in Fig. A.5.)

$$P'(\bar{x}, \bar{y}, \bar{z}) = \mathbf{A}_{\text{ref}} \mathbf{A}_{\text{out}}^{-1} P(\bar{x}, \bar{y}, \bar{z}). \tag{A.7}$$

The alignment transformation $\mathbf{A}_{\text{align}}$ (Sec. A.2.2) is used to transform the points between the tracker coordinate system and the image frame (first and last transformation steps in Fig. A.5.)

$$P'(x, y, z) = \mathbf{A}_{\text{align}} P'(\bar{x}, \bar{y}, \bar{z}) \tag{A.8}$$

$$P(x, y, z) = \mathbf{A}_{\text{align}} P(\bar{x}, \bar{y}, \bar{z}) \tag{A.9}$$

Inserting (A.8) and (A.9) into (A.7) gives the correction between a "new" point and the reference point in the image frame

$$P'(x, y, z) = \mathbf{A}_{\text{align}} \mathbf{A}_{\text{ref}} \mathbf{A}_{\text{out}}^{-1} \mathbf{A}_{\text{align}}^{-1} P(x, y, z), \tag{A.10}$$

and therefore also between a "new" acquired image (frames 2–11) and the reference image (frame 1). Thus the point transformation can be transferred to a transformation of a whole grid of points, assuming the patient tool is fixed to the skull and the brain is a rigid body. The corrected grid (reference grid) of the PET image is

$$\begin{bmatrix} x'_n & y'_n & z'_n & 1 \end{bmatrix}^{\text{T}} = \mathbf{A}_{\text{align}} \mathbf{A}_{\text{ref}} \mathbf{A}_{\text{out}}^{-1} \mathbf{A}_{\text{align}}^{-1} \begin{bmatrix} x_n & y_n & z_n & 1 \end{bmatrix}^{\text{T}} \tag{A.11}$$

where n is the voxel index of the PET image. Every coordinate in the rotated image is moved into the corrected grid (reference grid). Thus Fig. A.5 should

be read from right to left to get the corrected image matching the PET image coordinate system. The images were resampled to the discrete corrected grid of coordinates using trilinear interpolation.

Unfortunately, the Polaris system was not set to give the transformation of the patient tool relative to the reference tool during two of the scans (scan 7 and scan 8). Nevertheless all eight trackings were used. The transformations of the patient tool relative to the position of the reference tool were obtained in track 7 and track 8, similar to what is done in the Polaris system itself. The recorded inverse transformations of the reference tool were multiplied by the recorded transformations of the patient tool. The calculation of the motion correction is the same, irrespective of whether or not a reference tool is used, and the accuracy of the Polaris system should be the same.

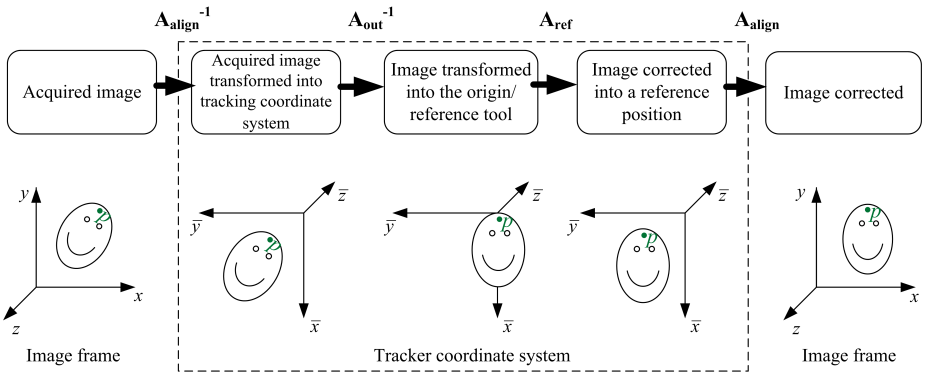


Figure A.5: Sketch of the repositioning of a reconstructed image to a reference position using external tracking information. The scheme involves two coordinate systems and four sub-transformations. The point P is the patient tool fixed to the subject.

A.2.5 Similarity Analysis of the Phantom

The effect of motion correction in the phantom experiments was quantified by calculating a percentage overlap (POL) [40, 155], i.e., a volume-based overlap between an initial reference frame image and a later non-corrected or corrected image. As the inner diameter of the tubes (1.5 mm) was approximately the same size as the spatial resolution in the HRRT images, the threshold value of the POL calculation was adjusted to obtain two continuous tubes on the images. The number of voxels used in the calculation was set to a value corresponding to a tube diameter equal to 2.5 times the measured inner diameter of the tubes.

It was observed that for tube diameters less than 2.5 times the actual measured diameter, the coiled tubes were not continuous on the images. The number of voxels representing the tubes was therefore calculated as

$$N = 2\pi(2.5r)^2 l, \tag{A.12}$$

where r and l are the radius and length of each tube, and N is the number of threshold voxels. The threshold value was set such that the number of voxels with intensity higher than the threshold value was equal to N, so each image includes N voxels representing the tubes. The threshold was estimated individually in each image using an iterative procedure.

The POL between two thresholded images a and b of the phantom was calculated as the number of overlapping voxels $v(a, b)$ relative to the estimated number of voxels within the tubes as follows

$$POL(a, b) = \frac{v(a, b)}{N}. \tag{A.13}$$

A.2.6 Simulation

Movements of the phantom were simulated to estimate the sources of errors and evaluate the POL as a quantitative measure. Additionally, a lookup table for converting the POL measure into an mm-displacement of the phantom was obtained by the simulation. One frame of the phantom was rotated around and translated along the principal axes in steps of either 1 degree or 1 mm up to 10 degrees rotation or 10 mm translation. At each position, the POL between the frame and the moved frame (non-corrected) or the inverse transformed frame (corrected) was calculated. Thus the corrected frames represent the forward transformation into a moved position, followed by the inverse transformation back to the initial reference position.

A.2.7 Cross Test

The measurements by the Polaris may not accurately represent the actual motion. To validate this possible lack of accuracy of the measurements and the repeatability of the phantom's motions, we conducted a cross test. The precision of the tracked motion and the precision of the performed motion were cross tested by calculating the POL of the scans using tracking information recorded during other scans. The frames with 10 degree rotation were considered for this test.

Data were expressed as mean ± standard deviation (SD). Two-sided ANOVA analysis was used to test significance within groups between factors.

A.2.8 Tracking Test

The tracked angle of the ten-degree rotation of the phantom was estimated in order to test the repeatability of the Polaris tracking (motivated by our preliminary cross test results). The recorded positions of the patient tool were fit to the rotation plane by a principal component analysis (PCA). The points were then fit to a circle from which the angle was estimated without further determination of the accuracy.

A.3 Results

Figure A.4 shows an image of the complete phantom and Fig. A.6 shows 3D PET EM images of the phantom before and after motion correction of rotation. The images are summed over frames 2–11 (1–10 degree rotations). Figure A.6(a) and Fig. A.6(b) show the sum image before and after correction respectively. The phantom in the reference position at 0 degrees (frame 1) is seen in Fig. A.6(c), and in Fig. A.6(d) the reference image is plotted on top of the corrected summed frames. The POLs between the reference image and the non-corrected and corrected images at 10 degrees (frame 11) are 15% and 76%, respectively.

A.3.1 Simulation

The simulation results for the misaligned phantom are seen in Figs. A.7(a) and A.7(b). The POL was calculated between a reference image of the phantom and the same image moved around or along the principal axes. The POL decreases linearly by approximately 20% per millimeter translation up to 4 mm displacement and approximately 13% per degree rotation up to 5 degrees.

The POL is reduced to 77–72% when the image of the phantom is misaligned with a reference position by as little as 1 mm. This should be noted for the later conversion of POLs into millimeter misalignment of the phantom in the next section. A misalignment of 0.5 mm corresponds to a POL of 86–89% as indicated by the blue lines in Fig. A.7(a).

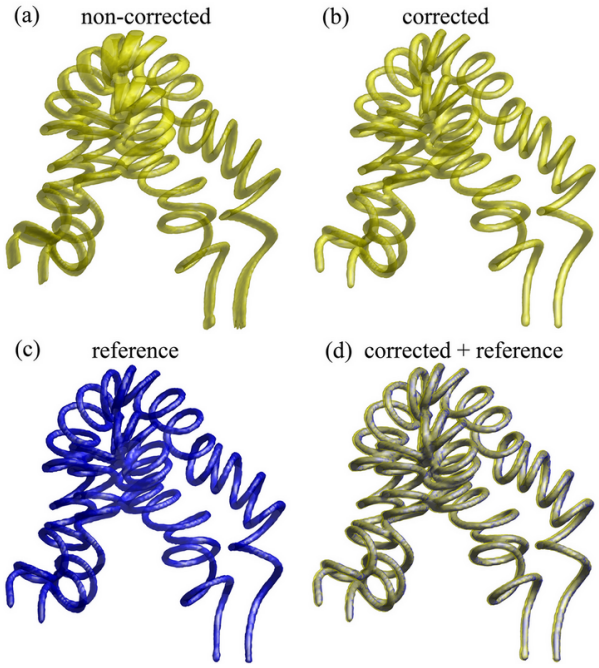


Figure A.6: 3D surface plot of the PET EM images of the phantom. (a) Summed images blurred over 10 degrees axial rotation. (b) Summed images at 1–10 degrees of the corrected frames. (c) Reference frame at 0 degrees. (d) The corrected summed images plotted on top of the reference frame.

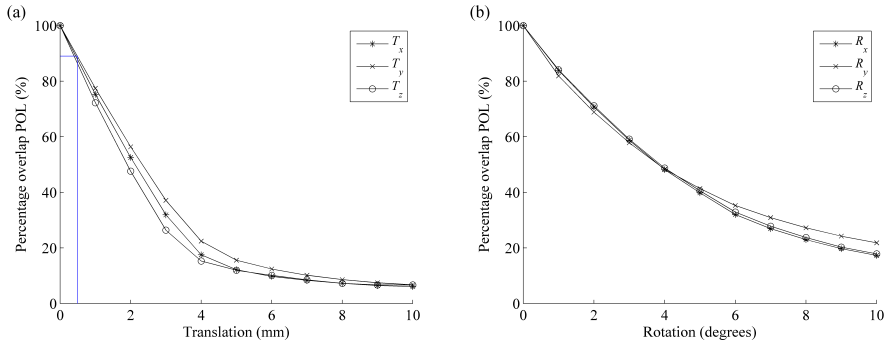


Figure A.7: POL as a function of simulated movements along (a) or around (b) the three principal axes (x: left/right, y: anterior/posterior, z: axial axis). The blue lines indicate an error of 11%, corresponding to a maximal misalignment of 0.5 mm of the phantom.

A.3.2 Phantom Experiments

The experimental results of the phantom moved by the rotary stage motor in steps of 1 degree are shown in Fig. A.8. The POL is shown as a function of the axial rotation before and after the images are repositioned into the reference positions at 0 degrees. The eight scans are shown as blue curves. The curves of the non-corrected images are the same within a POL of 2% and go down to 15.2% ±0.26% at 10 degrees rotation. This result proves the high repeatability of the motion of the phantom and shows the stability of the phantom design.

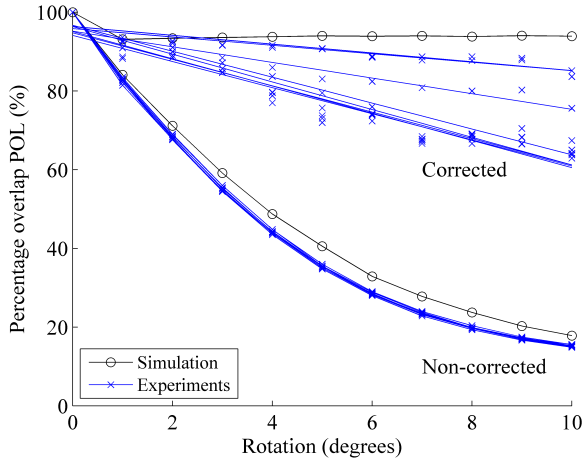


Figure A.8: POL as a function of axial rotation before and after the images are motion corrected. Blue curves represent the eight experiments. The two black curves represent simulation results where the lower one is the same as Rz in Fig. A.7(b), and the top one is the result of a forward motion transformation followed by the inverse transformation of a reference frame.

The POLs of the motion corrected images from the experimental studies are all significantly improved compared to the non-corrected images. The POLs of the corrected images decrease with motion magnitude. At 10 degrees the POLs of the corrected images are 63–85% (71% ±9.2%). Comparing this result with the simulated translation results (Fig. A.7(a)) a POL of 63–85% corresponds to an average misalignment of the phantom of less than 2 mm (1 mm $\sim$ 77–72%). The POLs of the corrected images are not 100% due to the statistics in the measurements, interpolation errors, tracking errors, and POL estimation errors.

The POLs of the simulated movements of an axial rotation are included in Fig. A.8 as they match the motion of the experiments. The sources of errors are

estimated by comparing the simulation of a non-corrected and corrected image respectively with the experimental results. The error due to noise is estimated from the non-corrected images (lower curves in Fig. A.8). The POLs of the non-corrected images are 1–5% lower than the simulated result, which represents the statistical difference and noise between the PET scans. This is because, in the simulation the POL is calculated between a frame and the same frame rotated, while in contrast in the experiments the POL is calculated between frames with different PET acquisitions. Thus the difference calculated between the two cases is the statistical difference and represents the noise of the PET images. The POL between the full 11-minute stationary scan and the images of the same scan divided into eleven one-minute frames was 96.6% ±0.13%. This result supports the finding that the error due to statistical differences is around 3.4%. The error due to the difference between scans (including small differences of the phantom's position) is not represented in this result as it was included in the above noise estimate of 1–5%.

The interpolation error of the motion correction method is represented by the POL of the simulated corrected images, which is constant at 93.7% ±0.31% for rotations > 1 degree, as seen in the top black curve of Fig. A.8. That we get a value of 93.7% and not 100% is due to the two interpolations of the simulation. In the simulation a frame is rotated forth and back resulting in two interpolations, whereas in the experiments there is only one interpolation per corrected movement. Thus the interpolation error within the experiments is assumed to be less than 6%.

Altogether the interpolation error, noise errors, and also the uncertainty of the POL estimation represent a POL of at most 11%, a value indicated by a blue horizontal line in Fig. A.7(a), corresponding to a misalignment of 0.5 mm of the phantom indicated by the blue vertical line in Fig. A.7(a). Hence, correction of movements less than 0.5 mm would not improve the POL.

A.3.3 Cross Test

The results of the cross test are listed in Table A.1 where the POL is shown for all possible combinations of scans and trackings. Table A.1 lists the POL for a corrected movement of 10 degrees with the scans along the rows and the applied trackings along the columns (Track 1 represents the tracking recorded during scan 1 and so forth). The variation within a column represents the variation between the scans, including the performed motion, noise, and interpolation error. The variation along a row represents the variation of the registered motion. The variation between trackings (last column) is approximately seven times higher than the variation of the bottom row and this shows that the stage is more

consistent in placing the phantom, compared to actual Polaris measurements. The POL is significantly different when a tracking from another scan is used (Two-sided ANOVA, $p < 10^{-15}$) and moreover the highest POL is not the POL using the tracking recorded during the actual scan (POL of the diagonal) except for scan 8 (row 8). For all experiments the best result is obtained using track 8 in the frame repositioning and similarly the poorest result is obtained using track 3 (Table A.1).

CROSS TEST OF THE PHANTOM STUDIES BASED ON THE PERCENTAGE OVERLAP

		Trackings								
		Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8	μ ±SD
Scans	1	**64**	65	64	65	70	78	87	89	73 ±10.5
	2	65	**65**	64	66	69	77	87	89	73 ±10.4
	3	63	64	**63**	65	69	77	86	87	72 ±10.2
	4	63	63	62	**64**	70	78	84	86	71 ±10.2
	5	61	61	60	62	**67**	77	84	86	70 ±10.6
	6	62	63	62	64	67	**76**	86	88	70 ± 9.8
	7	60	60	59	61	68	77	**84**	86	70 ±11.3
	8	60	61	60	62	68	77	83	**85**	70 ±11.8
μ ±SD		62 ±1.7	63 ±1.7	62 ±1.8	64 ±1.7	69 ±1.2	77 ±0.7	85 ±1.5	87 ±1.4	

Table A.1: The table shows POL [%] for a corrected movement of 10 degrees rotation in the eight experiments. The columns list the POL of the different scans using the same tracking within the frame repositioning. The rows list the POL using different trackings within the frame repositioning recorded during the eight scans (Track 1–Track 8). The diagonal shows the POL using the tracking recorded during the actual scanning. The last column lists the mean (μ) ±SD [%] of the POL using the same tracking on different scans. The last row lists the mean (μ) ±SD [%] of the POL using different trackings on the same scan. It is seen that the variation between trackings is almost seven times higher than the variation between scans.

A.3.4 Tracking Test

Table A.2 lists the tracked motion during the scans at the 10 degrees position. The tracked motions were measured to 9.3–10.6 degrees, which is within 1–7% of the performed motion. This suggests that the relatively high variation of the POL of the corrected images (as seen in Fig. A.8 and from the cross test in Table A.1) could be explained by the low accuracy of the tracking.

RECORDED MOTION DURING EIGHT SCANS OF THE PHANTOM

Tracking	1	2	3	4	5	6	7	8
Motion	10.6	10.5	10.4	10.5	9.9	9.8	9.5	9.3
Deviation	6.1	4.5	4.0	4.9	1.4	2.3	4.7	7.0

Table A.2: Estimation of the tracked motion of ten degrees during the eight rotation scans of the phantom. Motion = tracked motion in degrees. Deviation = percentage deviation of the performed motion of 10 degrees rotation.

A.4 Discussion

The resolution of PET scanners for clinical use is improving rapidly towards the 1–2 mm resolution achievable with the HRRT research brain scanner. This development makes the issue of brain motion correction relevant, not only for the more exotic high resolution research tomographs, but also for clinical PET scanners. A robust and practical approach to motion correction is therefore relevant.

This study describes a new moveable phantom that can be used for testing motion corrections in data sets acquired from modern high resolution PET scanners. The phantom was tested applying frame repositioning motion correction using a Polaris Vicra system as an external tracking system for measuring motions. The new phantom was developed to simulate a patient's head movements during scanning. A novel design with coiled tubes filled with radiotracer was used to represent the human brain. So far, only non-moveable phantoms have been used and, in contrast to the new mechanical phantom described here, they cannot simulate long drift motions or make accurately repeatable movements. Our phantom can perform computer-controlled movements and has a stable mechanical design. The phantom can simulate slow continuous movements and rapid gradual movements around and along the axial axis during scans.

The design of the phantom was stable. Simulation of axial rotation of a complex 3D structure was highly reproducible. The percentage overlap of eight scans performed over three days was within 2% before the images were corrected for motions (Fig. A.8).

One of the aims of the phantom was to be able to make reproducible motions, for example to evaluate sources of errors of a correction method. Assuming the performed motion is exactly the same for all scans, the variation along the rows in the cross test (Table A.1) is only due to the uncertainty of the tracking system. The assumption is that the repeated motions can also be validated

from the cross test itself by looking at the low variation of any one of the trackings applied to the different scans (SD of the last row in Table A.1). This measured SD of 1–2% was expected due to noise and interpolation errors. Thus the performed motion is highly repeatable. The SD of the POL between the trackings was 10−12% (last column of Table A.1), that is seven times higher than the SD of the POL between the scans, demonstrating that the high variation of POLs of the corrected images (Fig. A.8 blue top curves) is mainly due to the uncertainty of the tracking. The results of the recorded rotations during the eight scans (Table A.2) support the notion that the tracking is a significant source of error. The recorded rotations deviated by 1–7% of the performed rotation of ten degrees. Two of the trackings (Track 7 and Track 8) were not recorded automatically relative to the reference tool. They were calculated in the same way as in the Polaris software after recording, but they could potentially be slightly different. An obvious speculation is whether it may be better not to use a reference tool, since the POLs of scan 7 and scan 8 were the best results (Table A.1). However an additional scan 0 that was excluded from the results due to different activities in the two tubes of the phantom, gave POLs of 79–85% (track 0 used on scans 1–8) *with* the reference tool being used. Therefore we cannot conclude that tracking *without* the reference tool gives better results. It seems that the POL improves with track number, independent of scans when looking at Table A.1. Since the first track recorded, track 0, gave better POLs than tracks 1–6, we find that there is no correlation between tracking performance and time.

We have introduced a volume-based analysis method for the phantom studies. The sensitivity of the method was demonstrated to be high in all directions as the POL decreases, with approximately 20% per mm and approximately 13% per degree (up to 4 mm and 5 degrees, Fig. A.7(a) and Fig. A.7(b)). The high sensitivity, especially to small movements, is very useful in at least three ways: a) for comparing the performance of different correction methods, b) when evaluating sources of errors of a correction method, and c) in evaluating which kinds of movements to correct for.

The MAF correction method has been demonstrated to work well on an ECAT Exact HR$^+$ PET scanner with a Polaris Vicra tracking system using both a 3D Hoffman brain phantom and patient data [64]. That study showed the visual effect of the MAF correction, but here we present a quantitative analysis on a high resolution scanner. The quantitative analysis of the MAF motion correction on the HRRT with our phantom design using the Polaris Vicra system showed that the frame repositioning motion correction method improves the image quality when correcting for rotations of 1 degree or higher (Fig. A.8 blue curves). The POL was improved in all cases, yielding a misalignment of less than 2 mm of the corrected images, of which approximately 0.5 mm was due to noise and interpolation errors.

The results of the motion correction were dependent on the accuracy of the registered motions, which was demonstrated by the reduced variation of POLs using the same tracking output on all the scans (using any of the tracking outputs as seen in Table A.1). The POL was correlated with the tracking independently of the scan. The best results were the POLs obtained applying track 8 and the poorest results were the ones applying track 3. It was expected that the best results would be the diagonal of Table A.1, where the POLs were based on trackings recorded during the actual scans. Our results indicate that the variation is mainly caused by the uncertainty of the Polaris tracking and support the repeatability of the phantom setup. The tracking tool was fixed on the phantom with different orientations over the scans and the tracking accuracy differed within the tracking volume, which explains the variation of our results. This signifies that the registered head movements vary depending on where the tool is located on the head.

Our study quantified the effect of the sources of errors on the motion correction results and the accuracy of the tracking system. The repeatability and sensitivity of the phantom design showed that the POLs of the non-corrected images were within a range of 2% for movements of 1–10 degrees (Fig. A.8), thus the phantom design and POL analysis are useful for quantitative studies of movement artifacts. The phantom can simulate fast and slow movements of different sizes with high repeatability. Repeatability is necessary to a) evaluate the effects of motion correction on different scan protocols, and b) to evaluate the effects of the different sources of errors in order to improve individual correction methods.

A.4.1 Limitations, Challenges and Future Perspectives

The images of the phantom were not corrected for scatter and attenuation because the phantom was designed with material of low attenuation. This was chosen to limit the sources of error within the reconstruction. However, this also means that the reconstruction and correction method will differ from scans on patients. The MAF method would include repositioning of the attenuation map before the EM scans were reconstructed. Therefore the threshold of when to correct for motions of the phantom may not directly be transferred to patient studies.

Our phantom simulates the majority of general patient motions, though the phantom is not a true reflection of e.g. the movements of a patient with Parkinson's disease. However, our phantom is a reliable tool for online 3D measurement of precisely known movements.

The choice of frame repositioning, sinogram, or event-based motion correction in PET have been generally discussed, however it has not yet been shown which correction scheme should be used for a given scan protocol or motion in PET brain imaging. In theory motion should be corrected for within the listmode data but in practice this is not trivial, and some assumptions will be needed, e.g. with respect to the normalization.

In terms of deciding when and how to correct for movement, one approach is to make a very intelligent thresholding depending on: a) number of movements within a frame, b) kind of movement, c) size of movement, d) type of scan protocol, and e) region of interest. The phantom will be used for such studies in the future.

A.5 Conclusion

A moveable phantom with a novel design was developed to simulate axial rotation and axial translation of the head during PET scanning. The phantom was stable and its setup and use with the HRRT was reproducible. The phantom is useful for measuring movement artifacts in high resolution scanners, here represented by the HRRT, and the design is a low-cost "do-it-yourself" model.

Acknowledgments

The authors would like to thank the John and Birthe Meyer Foundation which donated the HRRT scanner to Rigshospitalet. We would also like to thank the staff at The Mechanical Workshop at Rigshospitalet who helped us produce the mechanical devices.

Appendix B

Structured Light 3D Tracking System for Measuring Motions in PET Brain Imaging

Oline V. Olesen, Morten R. Jørgensen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen

Abstract

Patient motion during scanning deteriorates image quality, especially for high resolution PET scanners. A new proposal for a 3D head tracking system for motion correction in high resolution PET brain imaging is set up and demonstrated. A prototype tracking system based on structured light with a DLP projector and a CCD camera is set up on a model of the high resolution research tomograph (HRRT). Methods to reconstruct 3D point clouds of simple surfaces based on phase-shifting interferometry (PSI) are demonstrated. The projector and camera are calibrated using a simple stereo vision procedure where the projector is treated as a camera. Additionally, the surface reconstructions are corrected for the non-linear projector output prior to image capture. The results are convincing and a first step toward a fully automated tracking system for measuring head motions in PET imaging.

B.1 Introduction

Positron emission tomography (PET) imaging has changed the impact and standards of nuclear medicine. In the last decade PET image quality has improved considerably and comprehensive use of radiochemistry with new advanced PET technology has further developed the methodology. Patient motion lowers image quality, especially for high resolution PET scanners. The High Resolution Research Tomograph (HRRT, Siemens) is a brain-dedicated scanner, with a resolution down to 1.4 mm shown in [115] when reconstructing using a new 3D ordered subset expectation maximization (OSEM) reconstruction algorithm with resolution modeling. This method incorporates a spatially invariant point spread function (PSF) [26, 154]. If no corrections for patient head motion are performed, head motions lead to degradation of image quality especially during long acquisitions [58]. This degradation increases with increasing scanner resolution and thus head motions end up counteracting the technological advances of high resolution scanners.

Figure B.1: New tracking approach setup on the HRRT PET scanner at Rigshospitalet, Copenhagen, Denmark. The tracking system is mounted at the top of the HRRT gantry pointing at a volunteer's face.

Due to the low count rate and resulting low contrast information of the patient motion is assumed known for most of the suggested motion correction methods in PET imaging [39, 123, 132, 167]. An optical real-time motion tracking system has been preferred (Polaris System, Northern Digital Inc.) [93]. This system registers a rigid tracking tool with 3–6 infrared reflecting markers. The tracking tool is fixed to the Patient's head using different types of band-aid, helmets, wet-caps, or goggles. It has been reported that, using fixation, these methods can cause artifacts on the PET images [64]. Furthermore, experience shows that it is difficult to use markers in clinical settings. Experience also shows that marker based tracking system has problems registering the markers in the narrow scanner geometry and the accuracy of the system affects the motion correction results.

The purpose of our research is to develop a new 3D head tracking system that: (1) does not need any markers; (2) fits to the narrow geometry of the HRRT PET scanner; and (3) can potentially be built into future PET scanners. This paper focuses on describing a new tracking setup for PET brain imaging like the one shown in Fig. B.1.

An obvious choice would be a stereo vision system with two or more cameras operating fully automatically and does not need any markers. However these systems often use relative complex algorithms which are non-suitable for a real time tracking system. Moreover the spatial resolution of a stereo vision system without markers could potentially be improved using structured light. The problem in a stereo vision system is to find the correspondence between the image planes of the cameras, and this is simpler when using a projector. The projector generates feature points or lines in one image plane which are matched with e.g. lines on the captured images from one or more cameras in order to obtain correspondence.

Different approaches to projected patterns have been suggested in order to solve the correspondence problem. Generally, systems have preferred to use reconstruction methods based on line-shifting, where the correspondence between the projector and camera(s) is binary coded. This is referred to as Gray coding. Every line of the projector is coded with a given sequence of e.g. black and white colors within a number of projected patterns. This method is very robust; however, it is slow due to the large number of patterns needed. Phase-shifting interferometry (PSI) has become a competitive alternative with the new DLP projector technology [43]. The method utilizes the colors which are related to the frequencies of a cosine function, since this can be considered as an analogue function theoretically it is possible to achieve sub-pixel resolution. In contrast to Gray coding, PSI is a fast method where as few as three patterns are needed per surface reconstruction which can be captured in one image using the three RGB color channels in a color camera [73]. Our system is to be used for highly accurate real time tracking and therefore it is based on PSI. One problem of using the colors of the patterns is, of course, that the output of the projector is non-linear. We apply a correction on the projected patterns.

The main challenge of PSI is the phase unwrapping, meaning how to distinguish between discontinuities of the phase image. The modular phase with a 2π modulus results in discontinuities of the phase image. These should be removed in the phase unwrapping, whereas discontinuities that are due to edges and structure of the object should not be removed. For objects with simple surfaces, phase unwrapping using simple one-dimensional path integration works very well [76]. However it is not a robust method for objects with sharp edges. Several more robust methods in two dimensions have been suggested. Normally these methods are classified into two groups: path-following algorithms and

path-independent algorithms. In two dimensions the paths for unwrapping are multi dimensional and for path-following algorithms aspects such as cut and cost for a given path are considered [54]. The path-independent algorithms include for example information on the singularity point's neighborhood [52] or larger regions [63].

B.1.1 System Requirements

The tracking system has to satisfy a number of technical and clinical requirements: (1) The registration of the position must be estimated simultaneously, so the detected PET event, known as a line-of-response (LOR), can be repositioned before the PET image reconstruction; (2) The tracking volume has to cover the region of the possible head motion in the HRRT scanner; (3) The system has to fit the narrow geometry of the PET scanner; (4) The accuracy of the tracking system has to be much smaller than the spatial resolution of the PET scanner (in the order of a tenth of a millimeter) otherwise the motion correction will increase the noise instead of reducing it; (5) The system must not interfere with the PET acquisition; (6) The sample frequency has to be a least twice as high as the head motions to avoid aliasing, according to the Nyquist criterion. However, due to the relatively low count rate in PET, a tracking frequency of 10–20 Hz should be adequate [93]. The clinical requirements are at least as important as the technical requirements. To be a part of clinical routines the tracking system must be: (1) Simple to use where a full automatic system is preferable; (2) The tracking system must have an easy interface with the PET scanner; (3) It must also be robust and have a flexible design to be a part of the daily routine; (4) The system must be comfortable for the patients, since an uncomfortable patient will introduce motions which are counter-productive for both patient's well-being and the image quality; (5) Finally the hygiene requirements of hospital use have to be met.

Commercial surface scanners are available; however they are not compatible with high resolution PET brain scanners as a tracking system for the following reasons. They do not fit to the narrow scanner geometry, they are not fast enough as a tracking system, or they use a visible light source. A visible light source is uncomfortable for the patients and might introduce motion or in at worst the scanning may be interrupted.

B.2 Materials and Methods

The idea is to build a 3D surface scanner into the HRRT PET brain scanner and use it to track patient motion during the PET scans. The principle of the pipeline is as follows: at each tracking frame a point cloud of the patients' head is reconstructed in the tracker coordinate system and aligned to a reference scan. The scan is obtained using a camera and a DLP projector. The point clouds of the surfaces are obtained based on PSI as described in Section B.2.4. This section only deals with finding correspondence between the image planes of the camera and those of the projector. The system calibration parameters from Section B.2.3 are used as the basis for converting the point correspondence of the image planes found in Section B.2.4 into points in the world coordinate system as described in Section B.2.5.

The tracking system should return the rigid transformations that align the reconstructed surfaces to a reference scan. This is beyond the scope of this paper. Here we demonstrate a proof-of-concept for the tracking setup and show preliminary results of reconstructed point clouds.

B.2.1 System Setup

Recent progress in projector hardware has improved the accuracy of surface scanners using PSI. To obtain the sub-millimeter accuracy made possible by this method, we set up a vision system consisting of a DLP projector (DLP Pico Projector, Texas Instruments) with HVGA resolution (480×320) and a grayscale CCD camera (Chameleon, Point Grey Research) with a resolution of 1280×960. The system was set up on a 1:1 model of the gantry of the HRRT PET scanner (Fig. B.2(a)). The image plane of the DLP projector consists of small micro mirrors that are switched on and off to control every pixel of the projected image. The size of this micro mirror device (DMD) is 2.42 mm times 3.63 mm and the size of the CCD chip is 3.60 mm times 4.80 mm. The camera and projector are connected to a computer and synchronized through a custom software setup. Patterns are projected onto the face of a mannequin head and captured as images by the CCD. The region of interest is chosen to be between the eyes where the facial motions are limited as illustrated in Fig. B.2(b).

The system was adjusted on the HRRT model to fit the narrow scanner geometry. The system was optimized to have the camera positioned 10–20 cm from the subject, the distance between the projector lens and the camera lens was 7 cm, which results in an angle of 25 degrees between the image axes of the camera and the projector. A mounting plate and a moveable holder with the above

distances were made for future experiments on the HRRT scanner (Fig. B.1).

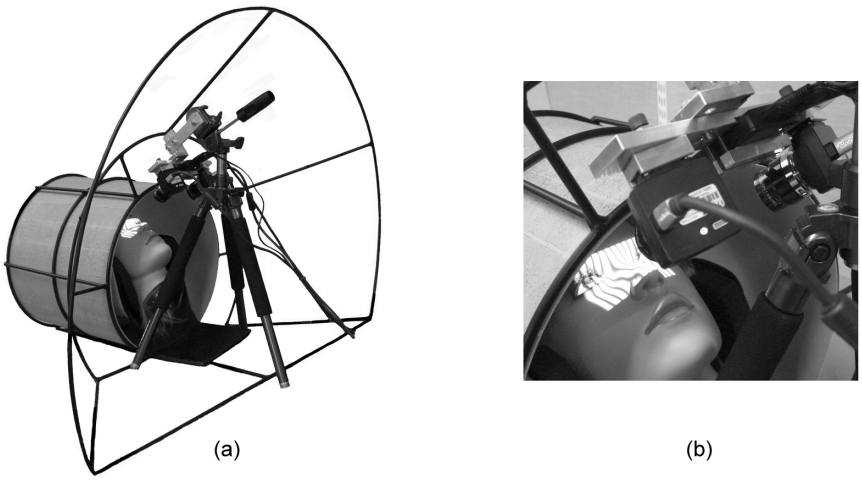


Figure B.2: (a) The prototype of the tracking system on the HRRT PET scanner model. (b) Zoom image of the system and a pattern projected onto the region of interest of the mannequin head. Two cameras are seen on the images but only one is used for the reconstructions in this study.

B.2.2 Experiments

Two sets of experiments were performed to achieve point clouds of surfaces including: (1) reconstructions of a mannequin head to demonstrate the method's usability on faces and the efficiency of correction for the non-linear projector output and (2) reconstructions of a plane to validate the scan accuracy. Four patterns were captured for each point cloud reconstruction, assuming the object was not moving during the capture. The four patterns were three shifted cosine patterns and a center-line pattern where the latter was used to generate an absolute phase-map.

To demonstrate the usability on objects with complicated geometry as the human face we used the mannequin head shown in Fig. B.2(b). The four patterns for the reconstruction were captured with and without correction of the non-linear projector to demonstrate the high impact on the reconstructions. The pitch p of the projected cosine patterns (corresponds to wavelength) was 40 pixels.

Finally, to demonstrate the quality of the scanning and its potential for tracking, reconstructions were obtained of the calibration plane. Four images of projected patterns for the phase-shift reconstruction were also captured at each calibration position to estimate the scan error beside the captured calibration images. Point clouds of the plane in eighteen calibration positions were reconstructed. The pitch of the cosine patterns was 20 pixels and the correction of the non-linear projector was applied on the projected patterns.

Evaluation of the scan error was based on root-mean-square errors (RMSE) to a reference. The RMSE of the point cloud to a fitted plane was estimated referred to as the relative RMSE. The planes were reconstructed where the extrinsic calibration parameters were set to this particular position, thus the plane should be at $z = 0$ in the world coordinate system. The RMSE was also estimated from this reconstructed point cloud to the z plane referred to as the absolute RMSE.

B.2.3 System Calibration

The components of the tracking system have to be calibrated to obtain points in the world coordinate system (x, y, z). The system setup is sketched in Fig. B.3 with two image planes, one representing the CCD chip (u_c, v_c), and one representing the DMD chip of the projector (u_p, v_p). Stereo calibration between the camera and the projector was obtained as a general stereo vision calibration with two cameras by creating images that look as if they have been captured by the projector (DMD image). The camera images of the calibration object were warped from the camera's CCD location to the projector's DMD location using a projected set of feature points on a calibration checkerboard. This method is similar to [172] though in our method we find correspondence using a projected grid of feature points whereas in [172] they use PSI as described in Section B.2.4. The calibration procedure was as follows.

Calibration images of a checkerboard with square sizes of 4 mm were captured in eighteen different positions *with* and *without* projected feature points on the checkerboard. Fig B.4(a) shows an image with feature points projected by the DLP. The checkerboard was white and gray coloured, and not black, so that it was possible to identify the feature points on the checkerboard background. Fig. B.4(b) shows a subtracted image at the same position *with* and *without* projected feature points. The feature points were identified automatically by blob detection indicated with the blue circles on Fig. B.2(b) [92]. The calibration images (CCD image) were then warped into the location of the DMD chip (DMD image) based on the coordinates of the projected feature points (red dots) and the identified feature points on the CCD images (blue circles), as indicated in

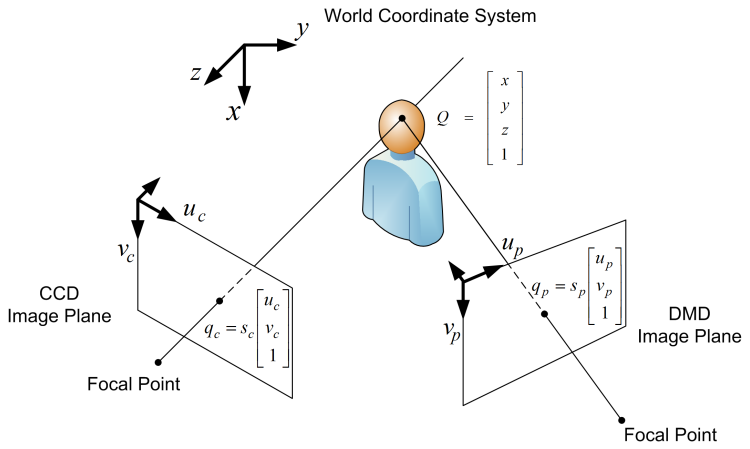


Figure B.3: Sketch of the relation between the world coordinate system (x, y, z) and the image planes for the camera (u_c, v_c) and the projector (u_p, v_p). The camera and the projector are both considered using the pinhole model. Q is a point in the world coordinate system and q_c and q_p are points in the CCD image plane and the DMD image plane respectively.

Fig. B.5 [61]. The intrinsic and extrinsic calibration parameters were obtained from the two sets of images: CCD images and DMD images at the eighteen positions, as generally done using [18] based on [173]. Two sets of calibration parameters were estimated based on nine positions each. The sets were obtained on two different days.

The white horizontal and vertical lines in Fig. B.5 are centerlines of the DMD image plane. These lines were also projected onto the checkerboard to validate the warped DMD images, which is seen correctly in the center of the DMD image in Fig. B.5(b).

The perspective camera model (pinhole model) was used for both the camera and the projector (Fig. B.3). The positions of the camera and projector in the coordinate system of the object *world coordinate system* (x, y, z) are described by the extrinsic calibration parameters $\mathbf{M}_c$ and $\mathbf{M}_p$ respectively

$$\mathbf{M}_c = [\mathbf{R}_c \mathbf{t}_c^{\mathrm{T}}] \quad \text{and} \quad \mathbf{M}_p = [\mathbf{R}_p \mathbf{t}_p^{\mathrm{T}}], \tag{B.1}$$

where $\mathbf{R}_c$ and $\mathbf{R}_p$ are the rotation matrices and $\mathbf{t}_c$ and $\mathbf{t}_p$ are the translation vectors. The complete models in homogeneous coordinates are

$$s_c \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \begin{bmatrix} f u_c & \alpha_c f u_c & \Delta u_c \\ 0 & f v_c & \Delta v_c \\ 0 & 0 & 1 \end{bmatrix} \left(\mathbf{R}_c \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) + \mathbf{t}_c \tag{B.2}$$

$$s_p \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} fu_p & \alpha_p fu_p & \Delta u_p \\ 0 & fv_p & \Delta v_p \\ 0 & 0 & 1 \end{bmatrix} \left(\mathbf{R}_p \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) + \mathbf{t}_p, \tag{B.3}$$

for the camera and the projector respectively. α is the skew coefficient (defining the angle between the u and v pixel axes), fu and fv are the focal lengths, Δu and Δv are the coordinates of the principal points in the image planes (u, v), and s is a scalar [60]. For simplicity, the coordinates in the images' planes are expressed as q_c and q_c, Q is a point in the world coordinate system ($Q = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^{\mathrm{T}}$), and $\mathbf{P}_c$ and $\mathbf{P}_p$ are the calibration parameters, including the extrinsic and intrinsic parameters. Thus the camera and projector models are

$$q_c = \mathbf{P}_c Q \quad \text{and} \quad q_p = \mathbf{P}_p Q. \tag{B.4}$$

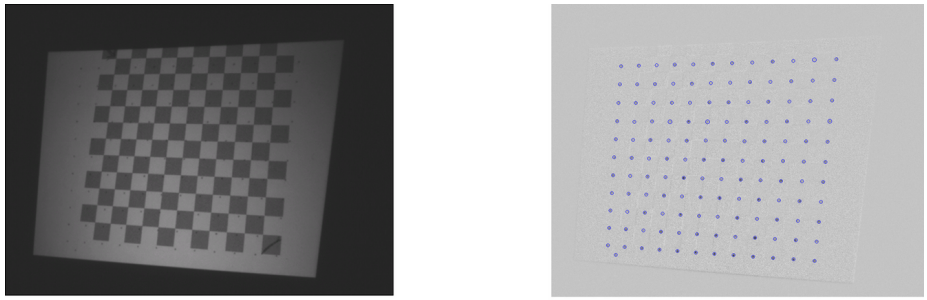


Figure B.4: (a) Captured calibration image of a checkerboard with projected feature points, nearly invisible to the human eye. (b) Subtracted image of two calibration image *with* and *without* projected feature points. The blue circles are indentified feature point using blob detection.

B.2.4 Phase-Shifting Interferometry

We use PSI to estimate the depth of a 3D object. This is a classic technique from the sixties [142]. The method utilizes the theory of wave propagation and interference of light. From a series of captured interferograms (2D images) the wavefront phase is computed and converted to line positions on the DMD image plane referred to as a phase-map. Thus the correspondence problem is solved by a phase-map where the pixel values of the CCD image plane (u_c, v_c) are associated with lines (u_p) on the DMD (see Fig. B.3).

In our case the wavefronts are generated as projected patterns. A series of three cosine patterns $I_k(u_p, v_p)$ are considered as wavefronts of two planar light beams

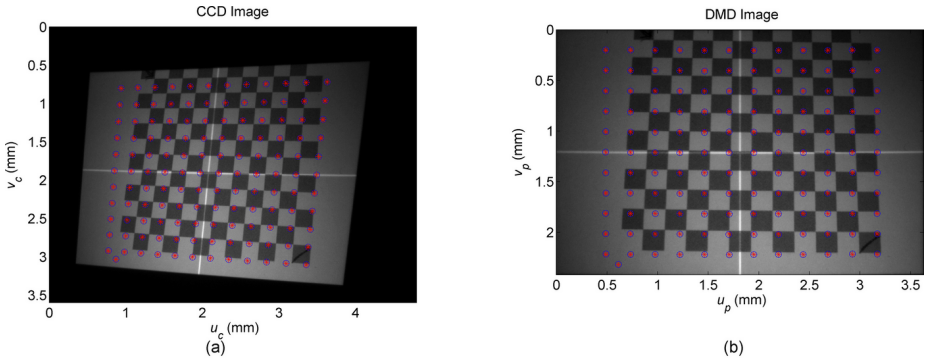


Figure B.5: Captured CCD image (a) and warped DMD image (b) with feature points based on the locations in the DMD chip (red dots) and the indentified ones on the CCD image (blue circles). The white lines are projected centerlines of the DMD image plane.

(a test beam and a reference beam) where the reference beam is varying in time. This time-shift is equivalent to a linear phase-shift s_k so the projected patterns can be expressed as

$$\begin{aligned} I_k(u_p, v_p) &= a\left[1 + \cos\left(\frac{2\pi}{p}u_p + s_k\right)\right] + b, \\ s_k &= \begin{bmatrix} \frac{2\pi}{3} & 0 & -\frac{2\pi}{3} \end{bmatrix}, \quad \text{and} \quad k = 1, 2, 3, \end{aligned} \tag{B.5}$$

where k is the pattern number, a is the amplitude, and b is the bias of the cosine function. The resulting interferogram or intensity pattern which is captured by the camera $cI_k(u_c, v_c)$ is then given by the fundamental PSI equation

$$cI_k(u_c, v_c) = I_{\text{av}} + I_{\text{mod}} \cos\left[\varphi(u_c, v_c) + s_k\right], \tag{B.6}$$

where $\varphi(u_c, v_c)$ is the phase, I_{av} is the average of the intensity, I_{mod} and is the intensity modulation. The phase can be computed without knowing the average intensity and the intensity modulation by the three captured images. Solving the equations given by (B.6) simultaneously gives the phase $\varphi(u_c, v_c)$

$$\varphi(u_c, v_c) = \tan^{-1}\left(\sqrt{3}\frac{cI_1(u_c, v_c) - cI_3(u_c, v_c)}{2cI_2(u_c, v_c) - cI_1(u_c, v_c) - cI_3(u_c, v_c)}\right). \tag{B.7}$$

This phase is more precisely the phase difference between the hypothetical test beam and the reference beam generated by the projector. The computed phase image (B.7) contains pixel values between 0 and 2π. The phase is modular due to the periodic cosine-pattern. Therefore the phase has to be unwrapped so that a given phase corresponds to a given line of the DMD image plane. Figure B.6 shows an example of a phase image and its unwrapped phase.

Several methods to perform phase unwrapping exist. A classic and very simple method is based on path integration of each line. First, discontinuities are found along the lines by edge detection and then a multiple of 2π is added or subtracted to the particular line interval. Two dimensional methods are more robust and can be classified into either path-following algorithms or path-independent algorithms. Experiments showed that the method described in [63] perform well with our data. This method is a two-dimensional path-independent algorithm where the image is divided into regions on the basis of the 2π phase jumps.

The captured image of the centerlines (center-line pattern) is used to identify the period/Fringe in the middle of the DMD image plane from which the phase is scaled to fit the horizontal size of the DMD. Thus the phase is converted to represent a vertical location u_p on the DMD.

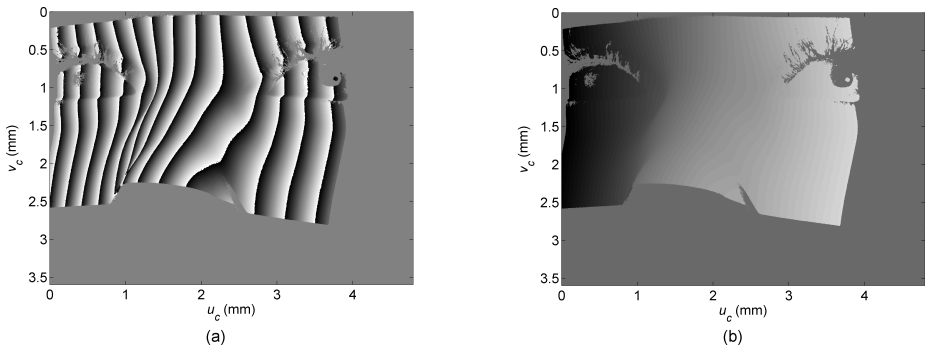


Figure B.6: Phase wrapped image (a) and phase unwrapped image (b) of the mannequin head shown in Fig. B.2(b).

B.2.5 Conversion to World Coordinates

The identified positions of the CCD and DMD image planes can be converted into points in the world coordinate system applying the system calibration parameters obtained from section B.2.3 [172]. The calibration matrices of the camera $\mathbf{P}_c$ and the projector $\mathbf{P}_p$ are 3×4 matrices and from the perspective camera model we have

$$q_c = \mathbf{P}_c Q \quad \text{or} \quad s \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \mathbf{P}_c \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{B.8}$$

Extracting the first or the second equations together with the last equation of (B.8) and isolating we get

$$u_c = \frac{\mathbf{P}_c(1)Q}{\mathbf{P}_c(3)Q} \quad \text{and} \quad v_c = \frac{\mathbf{P}_c(2)Q}{\mathbf{P}_c(3)Q}, \tag{B.9}$$

where the number of the calibration matrix represents a row e.g. $\mathbf{P}_c(3)$ is the third row of $\mathbf{P}_c$. Similar equations are valid for the projector. A new set of equations can be set up and solved with respect to coordinates in the world coordinate system. The coordinates in the CCD image plane and the vertical coordinate of the DMD image plane are expressed in three equations

$$u_c \cdot \mathbf{P}_c(3) - \mathbf{P}_c(1) \cdot Q = 0 \tag{B.10}$$
$$v_c \cdot \mathbf{P}_c(3) - \mathbf{P}_c(2) \cdot Q = 0 \tag{B.11}$$
$$u_c \cdot \mathbf{P}_p(3) - \mathbf{P}_p(1) \cdot Q = 0 \tag{B.12}$$

The new set of linear equations is then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{A}^{-1}\mathbf{b} \tag{B.13}$$

where $\mathbf{A}$ is a matrix and $\mathbf{b}$ is a vector consisting of calibration parameters.

B.2.6 2.6 Grayscale Correction

The output of the projector is non-linear. The colors in the dark and bright regions are flattened out, so the projector output corresponds to a gamma curve. In the phase-shifting reconstruction it is assumed that the captured images represent wavefronts of light with planar wave propagations. Thus the intensities represent true cosine patterns and therefore the colors are assumed linear.

As a consequence the patterns were corrected for non-linearity of the projector output prior to projection onto the object. A model of the projector output (gamma curve) was mapped to a linear scale and the inverse map was then applied on the patterns before they were projected.

B.3 Results and Discussions

B.3.1 System Calibration

The intrinsic calibration parameters were estimated on the basis of data obtained on two different days. The estimated errors in u and v of the image planes were 0.001 mm and 0.003 mm (SD) for the camera and the projector respectively (pixel error of [0.24 0.30] and [1.3 1.3] respectively).

The translations of the extrinsic parameter matched the psychical location of the camera and projector. The x-coordinates of the camera and the projector positions are measured to be 3 mm. Psychically, the projector lens was 3 cm above the camera lens (x direction) (the projector was turned upside down). The reason for this difference between the measured and observed distances is that the projector is designed to show images above the lens so it e.g. can be placed on a table without projecting onto the table itself. The measured difference in x was close to zero because we adjusted the visualized image planes of the CCD and the DMD to be similar at a distance of 15 cm from the system.

B.3.2 Mannequin Head

Figure B.7 shows images of the region corresponding to the one shown in Fig. B.2(b). The images are point clouds of the mannequin head's surface reconstructed *without* (left) and *with* (right) correction of the non-linear projector. The top and bottom images are the same surface seen from two different views. The ripples on the left images are artifacts due to the gamma-curve output of the projector. The images to the right in Fig. B.7 are the same image with grayscale correction and the ripples are virtually eliminated yielding a smooth surface consistent with the scanned mannequin. The captured images were thresholded to a value of 5% above the background. Thus dark regions as the eyebrows and shadows are excluded and seen as white regions on the images. Some of the exclude regions are due to occlusion which for example occurred on the lower left part of the nose. Occlusion can be limited by expanding the system to two cameras which also could improve the resolution further. Each point on the surface represents an observation, and the seemingly smooth surface is a result of a very low noise level of the system. This result is promising for accurate tracking of the position between frames.

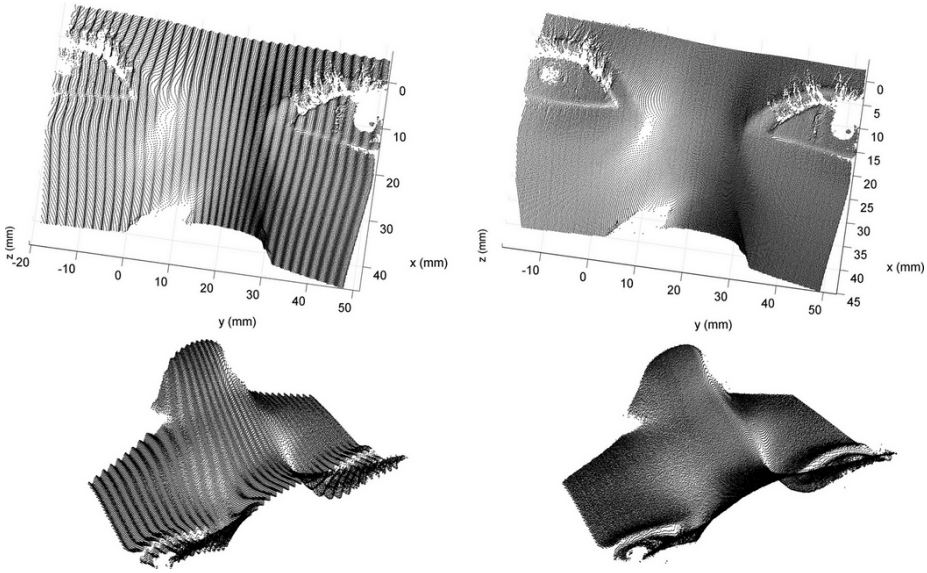


Figure B.7: Point clouds of the mannequin head's surface *without* (left) and *with* (right) correction of the projector output. Top and bottom images are of the same point cloud from two different views.

B.3.3 Plane

Figure B.8(a) shows an image of the eighteen reconstructed point clouds of the plane at the different calibration positions. The sizes of the planes are approximately 5×6 cm. The absolute RMSE and the relative RMSE of the planes are seen as a function of the plane number in Fig. B.8(b) as the bold curves. The means of the absolute and the relative errors are 0.40 mm and 0.06 mm respectively. Zhang and Huang made similar experiments in 2006 which are visualized in Fig. B.8(b) as the thin curve and seen to be between 0.10–0.22 mm [172]. That is 2.4 times higher than our results with respect to the mean results. This result indicates improved accuracy of the scanner when considering a relatively small measuring volume. It could be explained by the differences in the hardware and the methods used. The main difference compared to our method is the way the DMD images are created. Zhang and Huang created the DMD images using correspondences between the CCD and DMD image planes based on PSI, whereas we used feature points. Moreover the size of the measuring volume is important for the accuracy. Our system focuses on a small region due to motion seen in PET brain imaging of a few centimeters and thereby gains a higher resolution of the final tracking system. It should be noted that Zhang and Huang evaluated a larger measuring volume. They used

a plane with a size of approximately 16×23 cm.

B.3.4 Challenges and Future Perspectives

The system was demonstrated on a mannequin head and not on human faces. There are some challenges which have to be overcome in order to use it on patients, including: facial motions, occlusions, and the interaction of light with the human skin. A system with a visible light source is uncomfortable for the patient and not suitable for face tracking. Thus the system should be modified to use an invisible light source.

Further to using the system for tracking, it is obvious that the pose of the frames has to be computed and the time has to be optimized. The challenges are affordable and this preliminary study shows promising results for further development of a tracking system.

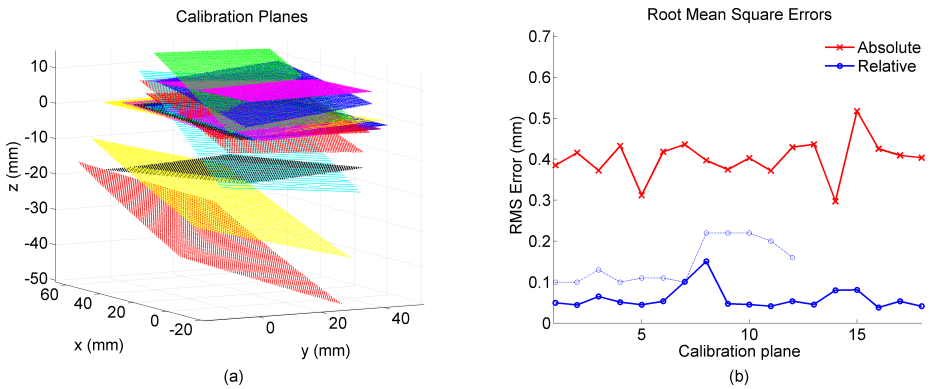


Figure B.8: (a) Point clouds of reconstructed planes at the eighteen calibration positions. The planes shown on the image are reconstructed using the extrinsic parameters of the position of the top black plane where $z = 0$. (b) The absolute RMSE and relative RMSE as a function of the number of the plane. The relative error is the RMS distance to a fitted plane and the absolute error is the RMS distance to the z plane where the plane is reconstructed using the actual position for the extrinsic parameters. The bold curves are the same planes. The thin curve is adapted results from Zhang and Huang of similar estimations of the relative RMSE, 2006 [172]. Thus this curve represents 12 other planes.

B.4 Conclusion

We have presented a first step toward an automatic and robust tracking system for high resolution PET brain imaging. A 3D head tracking prototype was set up and tested on a model of a high resolution PET scanner. The system was positioned and adjusted to fit the narrow geometry of the scanner. A state-of-the-art surface scanning technique was applied for reconstruction of 3D point clouds and demonstrated on the setup. Results of the scanner accuracy were shown to be accurate and reliable.

Acknowledgments

The authors would like to thank the staff at The Mechanical Workshop at Rigshospitalet, Copenhagen, Denmark and K.V. Olesen who helped us produce the mechanical devices. Ms. Olesen also thank the Danish Industrial PhD programme for supporting this research.

Appendix C

Motion Tracking for Medical Imaging: A Non-Visible Structured Light Tracking Approach

Oline V. Olesen, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen

Abstract

We present a system for head motion tracking in 3D brain imaging. The system is based on facial surface reconstruction and tracking using a structured light (SL) scanning principle. The system is designed to fit into narrow 3D medical scanner geometries limiting the field-of-view. It is tested in a clinical setting on the high resolution research tomograph (HRRT), Siemens PET scanner with a head phantom and volunteers. The SL system is compared to a commercial optical tracking system, the Polaris Vicra system, from NDI based on translatory and rotary ground truth motions of the head phantom. The accuracy of the systems was similar, with root-mean-square (RMS) errors of 0.09° for ±20° axial rotations, and RMS errors of 0.26 mm for ±25 mm translations. Tests were made using 1) a light emitting diode (LED) based miniaturized video projector, the Pico projector from Texas Instruments, and 2) a customized version of this projector replacing a visible light LED with a 850 nm near infrared LED. The

latter system does not provide additional discomfort by visible light projection into the patient's eyes. The main advantage over existing head motion tracking devices, including the Polaris Vicra system, is that it is not necessary to place markers on the patient. This provides a simpler workflow and eliminates uncertainties related to marker attachment and stability. We show proof-of-concept of a markerless tracking system especially designed for clinical use with promising results.

C.1 Introduction

The tomographic reconstruction of 3D and time varying 3D medical images from a series of scanning modalities including X-Ray, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) requires sequential data recording over time. Patient motion during acquisition time will result in a lower image quality or even render the examination useless for PET imaging [58]. Our focus is on the Siemens High Resolution Research Tomograph (HRRT) PET scanner, which is a brain scanner with an isotropic spatial resolution of 1.4 mm [115]. Motion induced image degradation increases with increasing scanner resolution and thus head motion counteracts the technological advances of high resolution scanners. The probability of patient motion occurring increases with acquisition time. For structural or anatomical imagery, patient motion can sometimes be estimated and compensated directly from the scan recordings, e.g. in cardiac MRI [152] and lung CT [36]. For functional 3D scans such as PET and fMRI low contrast and spatially sparse events hampers the direct estimation of motion from the recordings themselves, and external motion detection and tracking is preferred [123, 129, 132, 167]. An optical real time motion tracking system (Polaris System, Northern Digital Inc.) has been preferred for human studies [93].

An alternative optical tracker demonstrated on animals includes [53,164]. These systems register 3–6 infrared retro reflecting markers fixed in a position relative to each other. They are either glued onto the animal directly or mounted onto a tracking tool and then fixed to the subject. For human head tracking different types of band-aid, helmets, wet-caps, or goggles have been implemented. In a clinical setting attaching markers is an additional element in a busy work flow, additional discomfort to the patient, and with some of the fixation solution listed above the markers are likely to move independently of the patient's head with head motion giving rise to erroneous motion estimates. Experience also shows that a marker based tracking system has problems registering the markers in the narrow scanner geometry and the accuracy of the system affects the motion correction results.

Figure C.1: Photographs of a volunteer outside (top) and inside (bottom) the HRRT PET scanner with the SL system in the front mounted on the gantry.

The purpose of our research is to develop a new 3D head tracking system that (1) does not need any markers; (2) fits to the narrow geometry of the Siemens HRRT PET scanner; (3) is comfortable for the patients; and (4) can potentially be built into future PET scanners. Instead of tracking a geometrical object attached to the patient's head we propose to track the face itself. The human face has a rich collection of texture and color variations including common features, e.g. eyes, eyebrows, mouth, lips, and person specific features, e.g. moles. However, as we want to infer brain motion we concentrate on geometrical features that are less variant to facial expression, e.g. nose tip, the bridge of the nose, and cheekbones. These features are almost featureless in color and texture space and are characterized by their surface curvature and joint organization. A stereo vision system using structured light is an ideal way of recovering the surface of such color and texture less structures. By projecting a pattern onto the surface of interest we obtain surface features that when observed in a camera given proper geometrical calibration can be used to recover 3D surface point clouds. Different structured light patterns have been suggested including binary patterns, gray codes, phase shift gray scale patterns, stripe and grid indexing, and hybrid methods. [48] provides a comprehensive survey. A competitive tech-

nique providing a balance between the number of projections required and the complexity of the computations as well as providing subpixel resolution is the phase-shifting interferometry (PSI) method [43].

In order to track motion the recovered 3D point clouds must be aligned to a reference position. The limited system field-of-view (FOV) composed of: the camera projection FOV, facial geometry, and the size of rotation and translation will result in a limited overlap between the acquired point clouds. Thus a resistant and robust iterative closest point (ICP) alignment to an initial pre-computed template surface can be used to estimate rigid body motion of the head inside the scanner tunnel. The classical ICP algorithm [13] registers and aligns two point clouds with no prior correspondence. Several efficient variants of the ICP algorithm have been published [139].

In [111] we have shown first results using such a SL system based on projection of visible light patterns. In Fig. C.1 this prototype system is shown mounted on the Siemens HRRT scanner with a volunteer outside and inside the scanner. For the purpose of illustration the system is shown with the visible light projector such that the projected stripe pattern can be seen on the face of the volunteer. The system consists of a central video projector and two laterally placed cameras. The scanner imposes severe restrictions on the placement of cameras and projector and the subjects face is viewed at an oblique angle. While this system fulfills the requirements that it does not need markers, fits to the scanner geometry, and can potentially be built into future PET scanners it will also provide potential discomfort to the patient by shining light directly into the patients eyes.

In this system the illuminated area is chosen to be large enough to cover relevant parts of the facial geometry and small enough to primarily illuminate the face. Restrictions are at the same time imposed by the scanner geometry. The result can be seen in Fig. C.1. For large rotations and translations the overlap between the illuminated areas may be as low as 50%. Due to this partial overlap of point clouds a point rejection approach is applied with the ICP algorithm. Points matched to the border of the reference point cloud are rejected. This is non-trivial for raw point clouds. To overcome this, we represent our reference as a triangulated template surface computed using a state-of-the-art surface reconstruction algorithm [117]. Compared, to the popular Poisson surface reconstruction algorithm [82], the Markov Random Field surface reconstruction algorithm [117] deals particularly well with human body scans.

In this paper we present a new NIR-SL system based on a near infrared (NIR) video projector realized by customization of a Texas Instruments Pico projector. The system fulfills the requirement of not providing additional discomfort to the patient. We will review the PSI method and the point cloud alignment.

Moreover, we will provide quantitative evaluations of the performance of the SL and NIR-SL and compare these to the commercially available Polaris Vicra system. We demonstrate the system on a volunteer inside the HRRT PET scanner while not exposing the volunteer to any radioactivity.

C.2 System Requirements

The tracking system must satisfy a number of technical and clinical requirements: (1) The registration of the position must be estimated simultaneously so that a detected PET event known as a line-of-response (LOR) can be repositioned before the PET image reconstruction; (2) The tracking volume must cover the range of the possible head motion in the HRRT scanner; (3) The system must fit the narrow geometry of the PET scanner; (4) The accuracy of the tracking system has to be better than the spatial resolution of the PET scanner, otherwise the motion correction will increase the blurring instead of reducing it; (5) The system must not interfere with the PET acquisition; (6) The sample frequency has to be at least twice as high as the frequency of head motion to avoid aliasing, according to the Nyquist criterion. However, due to the relatively low count rate in PET, a tracking frequency of 5–10 Hz is adequate [93] or even less if applying the frame repositioning motion correction method [123]. The clinical requirements are at least as important as the technical requirements. To be a part of clinical routines the tracking system must be: (1) Simple to use with a preference for a fully automated system; (2) The tracking system must have an easy interface with the PET scanner; (3) It must be robust and have a flexible design to be a part of the daily routine; (4) The system must be comfortable for the patients, since an uncomfortable patient will introduce motion which is counterproductive for both the patient's well-being and the image quality; (5) Finally the hygiene requirements of hospital use have to be met.

Commercial surface scanners are available. However they are not compatible with high resolution PET brain scanners for the following reasons. They do not fit the narrow scanner geometry, are not fast enough as a tracking system, and do not use an invisible light source. A visible light source is uncomfortable for the patients and might introduce motion or at worst may interrupt the scanning.

C.3 Materials and Methods

We intend to integrate a 3D surface scanner into the HRRT PET brain scanner and use it to track patient motion during the PET scans. At each tracking frame

a partial 3D point cloud of the patient's head is processed in the tracker coordinate system and aligned to an initial template surface reconstruction referred to as the *reference target*.

C.3.1 Structured light tracking system

The SL system consists of a digital light processing (DLP) projector (DLP Pico Projector, Texas Instruments) with HVGA resolution (480×320) and two gray scale charge coupled device (CCD) cameras (Point Grey Research) with a resolution of 1288×964. The SL system is designed to match the narrow scanner geometry of the HRRT PET scanner with a flexible design that is easy to mount on the scanner gantry. The SL system mounted on the gantry of the HRRT PET scanner is shown in Fig. C.2 just above the patient tunnel. The image plane of the DLP projector consists of micro mirrors that are switched on and off to control every pixel of the projected image. This recent technology has improved the quality of the image projection and made it possible to achieve submillimeter accuracy of surface measurements [43]. The size of this digital micro mirror device (DMD) is 2.42 mm $\times$ 3.63 mm and the size of the CCD chip is 3.60 mm $\times$ 4.80 mm. The cameras and projector are connected to a computer and synchronized through a custom software setup. Patterns are projected onto the patient's face and captured as images by the CCDs. The region of interest (ROI) is around the bridge of the nose, seen as the bright region in Fig. C.1 bottom. This ROI is chosen due to limited facial motions at the bridge of the nose and because of the high surface curvatures. The system was optimized to have the camera positioned 10–20 cm from the subject and the distance between the projector lens and the camera lens was 8.5 cm resulting in an angle of 30 degrees between the image axis of the camera and the projector.

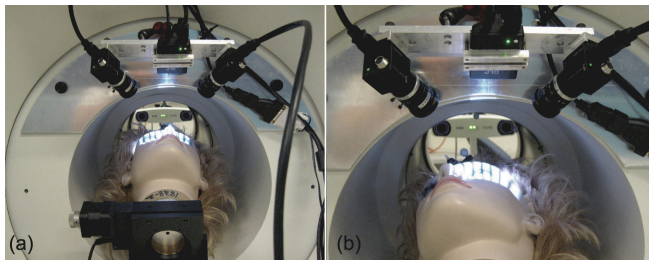


Figure C.2: Photographs of the mannequin head mounted to a rotary stage inside the HRRT PET scanner in two positions; (a) 0 degrees and (b) -20 degrees. The motor stage is seen in the bottom and the Polaris sensor in the back (the tracking tool can just be discerned above the forehead).

The SL system was modified with a NIR light emitting diode (LED) in order to meet the clinical requirements for such a system. The system has to be comfortable for the patients and shifting patterns of visible light projecting into the eyes of the patients is not acceptable during long acquisitions. A prototype of the NIR-SL system is seen in Fig. C.3 with a specially designed DLP projector (DLP Pico Projector, modified). One of the original three LEDs was bypassed by a NIR LED connected to a separate power supply allowing for adjustment of the light intensity. The inserted LED was cooled by an external fan fixed to the system.

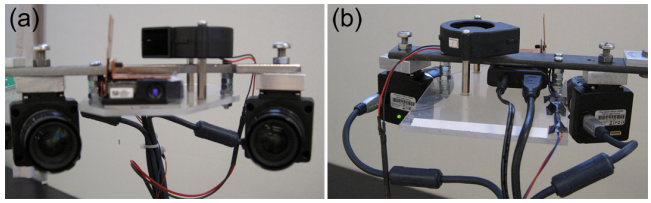


Figure C.3: The NIR-SL system with a modified Pico DLP. In the top a cooling fan is directed at the LED fixed to a copper plate.

C.3.2 Tracking experiments

1) Phantom study: The performance of the tracking approach was evaluated by a set of experiments on the HRRT PET scanner with simultaneous tracking using the Polaris Vicra system. A mannequin head was placed inside the patient tunnel as seen in Fig. C.2. It was mounted onto a nanorotary motor stage from Thorlabs. The stage made it possible to perform highly controllable rotations of the head. The stage was programmed to rotate in steps of 5 degrees from -20 to 20 degrees. The movements were repeated nine times. At each stationary position a set of four phase-shift images were captured with the SL systems. Furthermore, the phantom was translated in the axial direction in steps of 10 mm across 6 positions repeated four times. Axial translation is often seen when patients are moving their legs or relaxing neck and shoulders. The performed translation steps were measured with a sliding canvas with an estimated accuracy of ± 0.01 mm. At each experiment a 3D point cloud was reconstructed using PSI as described in Section C.3.4. This study extends the experiment described in [111].

We used the zero degrees position as the reference position and the pose of the head is estimated relative to this position. The alignment method has also been improved to handle greater motion compared to [111]. This was done by increasing the partial overlap between the reference target and the 3D point

clouds at each tracking frame. The reference target was based on 3D point clouds captured outside the HRRT PET scanner just before the subject entered the scanner. Thus a larger region of the face surface was represented in the reference target compared to [111] where point clouds from the initial position inside the HRRT PET scanner were reconstructed and used as the reference target. The reference target was moved into the position at zero degrees before the pose estimation of the 3D point clouds. The Polaris Vicra tracking tool was fixed to the forehead of the mannequin head as for patients to track the head motions during the PET acquisition [109]. While the Thorlabs stage provides baseline rotation data, the Polaris Vicra system recorded the motions of the head simultaneously with the image capturing of the SL system. Figure C.2 shows the set up of the experiments where the SL system is seen in the front and the Polaris sensor is seen in the back behind the patient tunnel. Figure C.2(a) shows the head in the reference position at zero degrees and Fig. C.2(b) shows the most extreme rotation of the head to the left at -20 degrees. The Polaris Vicra system directly provides a 3×3 rotation matrix $\mathbf{R}$ (with elements R_{ij}) and a translation vector $\mathbf{t}$ with respect to a reference position. The SL systems uses a software package *Sumatra* [118] for the ICP alignment returning $\mathbf{R}$ and $\mathbf{t}$. To be able to compare the rigid motion estimates from the SL systems and the Polaris Vicra system with the baseline translations and baseline rotations provided by the Thorlabs stage, the rotation angle θ, direction of rotation axis $\mathbf{v}$, and a point on the line $\mathbf{c}$ is determined from $\mathbf{R}$ and $\mathbf{t}$ [70]:

$$\begin{aligned} \theta &= \arccos\left((\text{trace}(\mathbf{R}) - 1)/2\right) \\ \mathbf{v} &= \frac{1}{2\sin(\theta)} \begin{bmatrix} R_{32} - R_{23} & R_{13} - R_{31} & R_{21} - R_{12} \end{bmatrix}^{\mathrm{T}} \\ \mathbf{c} &= (\mathbf{I} - \mathbf{R})^{-1}\mathbf{t}, \end{aligned} \tag{C.1}$$

where $\mathbf{I}$ is the identity matrix.

2) Human study: We tested the system on a volunteer inside the HRRT PET scanner combined with the Polaris Vicra tracking. The volunteer was positioned as a patient would be and fixed using a vacuum bag as is normally done in the clinical routine at Rigshospitalet, Copenhagen. A reference target was obtained outside the HRRT PET scanner just before the subject entered the PET scanner. The reference target was constructed by aligning 3D point clouds recorded in four different SL system positions and reconstructing a triangulated surface representation. In this way the reference target has full coverage of the face. Inside the HRRT PET scanner the volunteer was asked to move the head into fourteen different positions corresponding to movements often observed in PET brain imaging; (1) sidewise rotation, (2) upward rotation, and (3) axial translation. The registered motions are evaluated by comparing the angle of rotation relative to a mean position for each of the systems using Eq. (C.1). The mean

rotation $\underline{\mathbf{R}}$ is determined as [57]:

$$\begin{aligned}
\underline{\mathbf{R}} &= \arg\min_{\mathbf{R}} \sum_i \theta^2 \left(\mathbf{R}^{-1}\mathbf{R_i}\right) \\
&\approx \arg\min_{\mathbf{R}} \sum_i 3 - \operatorname{trace}\left(\mathbf{R}^{-1}\mathbf{R_i}\right) \\
&= \arg\max_{\mathbf{R}} \sum_i \operatorname{trace}\left(\mathbf{R}^{-1}\mathbf{R_i}\right) \\
&= \arg\max_{\mathbf{R}} \operatorname{trace}\left(\mathbf{R}^{-1}\sum_i \mathbf{R_i}\right)
\end{aligned} \tag{C.2}$$

The solution to (C.2) is found by singular value decomposition (SVD) $\sum_i \mathbf{R_i} = \mathbf{UDV}$. Introducing the matrix $\mathbf{S} = \operatorname{diag}(1, 1, \det(\mathbf{UV}))$ we have the mean rotation given as

$$\underline{\mathbf{R}} = \mathbf{USV} \tag{C.3}$$

C.3.3 Pose estimation

We wanted to estimate the rigid body transformation from the current 3D scan to the reference scan. The scans are unstructured point clouds where approximate estimates of the point normals exist. We are using a specialized version of the ICP algorithm [13]. Initially, two surfaces for each camera respectively are created based on 2–4 scan positions. Both cameras produce a 3D point cloud representation of the part of the head in its FOV. Scans for each camera are aligned and merged to create a reference target that covers the FOV of each camera using the method described in [118]. In this method the surfaces are created using the Markov Random Field surface reconstruction algorithm [117]. It is based on an implicit description of the surface combined with a regularization step that makes it well suited for human body scans. Since the surface reconstruction algorithm by default computes surfaces that extend beyond the point cloud, a postprocessing step is needed where the surface is cropped to fit the point cloud. This is done by removing parts of the surface that are not supported by reliable input points. Support is defined as being within a distance d of an input point, where d is estimated as the average neighbor distance in the input point cloud. The result is a polygonised surface patch, where the edge vertices are defined as having only one adjacent triangle. For each point in the current scan, the closest point on the triangulated surface is found using a kD-tree based approach. If the point falls on an edge vertex, the point match is discarded. The remaining point matches are used to compute the rigid body transformation using the solution found in [70]. Using this method the transformation bringing the current scan into alignment with the reference surface

is computed. Prior to the alignment noisy points, non connected points, and small isolated clusters of points were excluded from the point cloud, following the approach from [117]. The alignment of the partial face surfaces into the reference target is computed twice. In the first round, the point clouds are aligned to a target representing most of the face to generate a robust prealignment. In the second alignment round the target is reduced to include the stable part of the face reconstruction just around the nose bridge.

C.3.4 3D point cloud generation

We use PSI to determine the correspondence between the two image planes; the projector image plane (u_p, v_p) and the image plane of one of the cameras (u_c, v_c) (see Fig. C.4). From a series of three captured interferograms (2D images) the wave front phase is computed and converted to line positions on the projector image plane [73]. Thus a given phase of cosine patterns $I_k(u_p, v_p)$ on the captured images $cI_k(u_c, v_c)$ correspond to a position on the projector image plane after phase unwrapping. The cosine patterns are generated by:

$$I_k(u_p, v_p) = a\left(1 + \cos(\frac{2\pi}{p}u_p + s_k)\right) + b,$$

$$s_k = \frac{2\pi}{3}(k-2) \quad \text{and} \quad k = 1, 2, 3,$$

were a is the amplitude, b is the bias, s_k is the shift, and k is the pattern number of the cosine function. This results in the three captured interferograms:

$$cI_k(u_c, v_c) = I_{\text{av}} + I_{\text{mod}} \cos\left(\varphi(u_c, v_c) + s_k\right),$$

with the three unknowns; the phase $\varphi(u_c, v_c)$, the average of the intensity I_{av}, and the modulation of the intensity I_{mod}. Solving the above equation gives the phase:

$$\varphi(u_c, v_c) = \arctan\left(\sqrt{3}\frac{cI_1(u_c, v_c) - cI_3(u_c, v_c)}{2cI_2(u_c, v_c) - cI_1(u_c, v_c) - cI_3(u_c, v_c)}\right)$$

Since the phase is periodic, the phase has to be unwrapped to achieve a continuous phase image. Several methods to perform phase unwrapping exist. Experiments showed that the method described in [63] performs well with our data. This method is a two-dimensional path-independent algorithm where the image is divided into regions based on the 2π phase jumps.

The points on the image planes are converted into 3D coordinates using a simple pinhole model for both the cameras and the projector and assuming the calibrations parameters for all three components are known. The calibration matrices

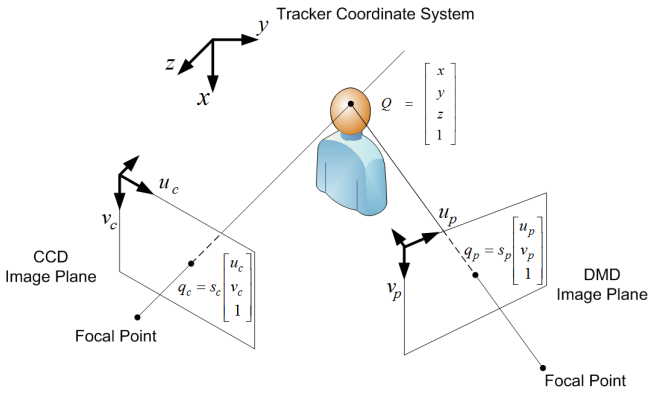


Figure C.4: Sketch of the relation between the world coordinate system (x, y, z) and the image planes for one camera (u_c, v_c) and the projector (u_p, v_p). Q is a point in the tracker coordinate system and q_c and q_p are points in the CCD image plane and the DMD image plane respectively.

of the cameras $\mathbf{P}_{c1}$ and $\mathbf{P}_{c2}$ and the projector $\mathbf{P}_p$ are 3×4 matrices and from the perspective camera model we have for one camera:

$$\begin{aligned} q_c &= \mathbf{P}_c Q \quad \text{or} \\ s[\ u_c \quad v_c \quad 1\ ]^{\mathrm{T}} &= \mathbf{P}_c[\ X \quad Y \quad Z \quad 1\ ]^{\mathrm{T}}. \end{aligned}$$

This can be combined into:

$$u_c = \frac{\mathbf{P}_c(1)Q}{\mathbf{P}_c(3)Q} \quad \text{and} \quad v_c = \frac{\mathbf{P}_c(2)Q}{\mathbf{P}_c(3)Q},$$

where the number of the calibration matrix represents a row e.g. $\mathbf{P_c}(3)$ is the third row of $\mathbf{P_c}$. Similar equations are valid for the projector. A new set of equations can be set up and solved with respect to coordinates in the tracker coordinate system. The coordinates in the CCD image plane and the vertical coordinate of the DMD:

$$\begin{aligned} u_c \cdot \mathbf{P}_c(3) - \mathbf{P}_c(1) \cdot Q &= 0 \\ v_c \cdot \mathbf{P}_c(3) - \mathbf{P}_c(2) \cdot Q &= 0 \\ u_p \cdot \mathbf{P}_p(3) - \mathbf{P}_p(1) \cdot Q &= 0 \end{aligned}$$

The new set of linear equations yields a 3D point in the tracker coordinate system:

$$s[\ X \quad Y \quad Z\ ]^{\mathrm{T}} = \mathbf{A}^{-1}\mathbf{b},$$

where $\mathbf{A}$ is a matrix and $\mathbf{b}$ is a vector consisting of calibration parameters. Further details of the system calibration and 3D coordinate computations can be found in [108].

C.4 Experimental Results

Figure C.5 (top) shows the 3D point clouds at the nine different positions from -20 degrees to 20 degrees (left to right) for one of the nine experiments with the mannequin head. The red and the blue point clouds represent the right and left camera respectively. As seen, the point clouds are highly detailed with little noise and outliers, demonstrating the high spatial resolution of the system. The right camera has a more favorable angular position with respect to the surface for the negative rotations, and the left camera for the positive rotations. Thus, we have used point clouds for the camera with the largest angle between the image axis and the surface as shown on Fig. C.5. The point clouds are aligned into the reference target and two results of the ICP alignment are shown in Fig. C.5 (bottom) at ±10 degrees from the right and left camera, respectively. In Fig. C.5 (bottom), the color coding of the aligned scans represents the individual per-point alignment error. It is computed as the distance from the point to the closest point on the reference target seen in the back. The errors between the reference target and the aligned points are in the order of 0–0.2 mm with the largest errors around the eyes. The medians of the point errors in Fig. C.5(bottom) are 0.10 mm and 0.09 mm.

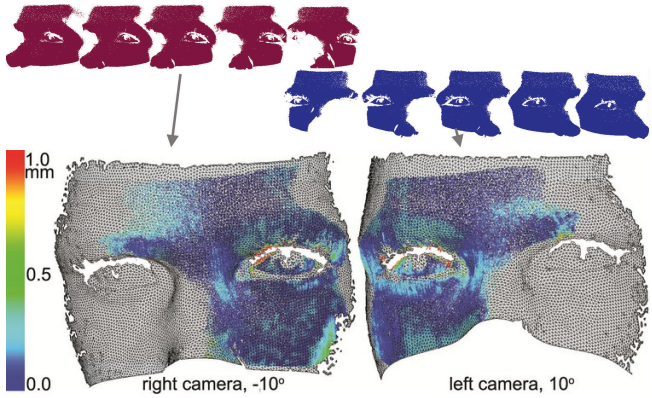


Figure C.5: Results of the mannequin head experiment. Top: 3D point clouds at the nine positions. Left to right from -20 degrees to 20 degrees. Blue images represent the left camera and red images represent the right camera. Bottom: ICP alignment at two positions ±10 degrees. Alignments into the reference target are shown on top of the reference surfaces. The colors represent the errors [mm] as the distance to target.

As previously mentioned, the motion of the Thorlabs stage is considered the ground truth motion. The errors of the estimated motion are plotted as a function of the ground truth motion in Fig. C.6. In Fig. C.6 (top) the results

from variation in the rotation is shown. The blue and red points represent estimates from the SL system's left and right camera, respectively. The black points are the simultaneously captured results from the Polaris Vicra system.

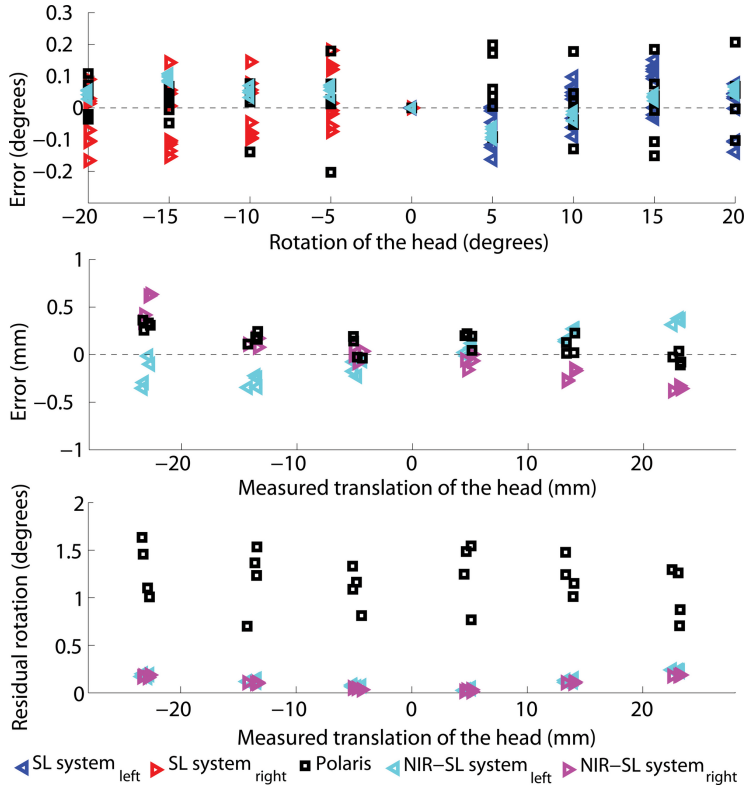


Figure C.6: Comparison of the SL system, NIR-SL system, and the Polaris Vicra system; Rotation study: (top) differences between the estimated and the performed rotations as a function of the performed rotation. Polaris Vicra rotation experiment 3 has errors > 0.5 degrees and not represented in the plot. Translation study: (middle) differences between the estimated and the performed translations as a function of the performed translation. (bottom) The estimated absolute rotations that ideally should be 0 degrees for the translation study.

The errors of the SL system are less than 0.2 degrees from the performed rotation when using the right camera (red) for negative rotations and the left camera (blue) for the positive rotation with a RMS error of 0.089 degrees. This is a similar result as the Polaris Vicra system, which has a RMS error of 0.086 degrees. One of the Polaris Vicra tracking (experiment no. 3) is treated as an

outlier and not included in the RMS error. We can not operate the NIR-SL system simultaneously with the Polaris Vicra system since both systems used light at 850 nm and would influence each others estimates. The NIR-SL systems has been used for the same experiment in a second run and the resulting errors are plotted in light blue. The NIR-SL system has a RMS of 0.061 degrees.

In Fig. C.6 (middle) the results from varying the translation is shown. Here we compare the NIR-SL system to the Polaris Vicra system. Again the experiment is done in two runs, one for the NIR-SL system, and one for the Polaris Vicra system in order to avoid interference. The light blue and magenta points represent estimates from the NIR-SL systems left and right camera, respectively. The black points are the results from the Polaris Vicra system. The RMS for NIR-SL (left) is 0.24 mm, for NIR-SL (right) is 0.28 mm, and for Polaris 0.18 mm. In Fig. 6 (bottom) the estimated absolute rotations for the translation study are shown. These should ideally be zero degrees. The light blue and magenta points represent estimates from the NIR-SL systems left and right camera, respectively. The black points are the results from the Polaris Vicra system. The NIR-SL system deviations are below 0.2 degrees. The Polaris Vicra ranges up to 1.7 degrees. The large deviation for the Polaris may be ascribed to the longer distance between the sensor and the object for this system.

Furthermore, we have tested the SL system on a volunteer in the HRRT PET scanner. The purpose is to demonstrate the clinical usability of the system. The position of the volunteer was registered in fourteen poses and estimated relative to the mean position. The subject had translated 5–43 mm and rotated 1–21 degrees compared to the mean position. Figure C.7 shows the results of the relative rotation in the fourteen recorded head poses and compares the three systems; the left SL system, the right SL system, and the Polaris Vicra system. The standard deviation (SD) between the differences of the systems are; (1) $\mathrm{SD}(\text{Pol-SL}_{\text{left}}) = 0.85^o$, (2) $\mathrm{SD}(\text{Pol-SL}_{\text{right}}) = 0.97^o$, and (3) $\mathrm{SD}(\text{SL}_{\text{left}}\text{-SL}_{\text{right}}) = 0.41^o$. It is seen that the difference between the systems increases with the angle of rotation from the centre, cf. Fig. C.7.

Figure C.8 shows the 3D point clouds before (left) and after (right) the ICP alignment to the reference target for two head poses. These are pose 7 and pose 10 which are close to the head poses shown in Fig. C.1. These poses are moved approximately 8 mm and 3 degrees for pose 7 and 35 mm and 19 degrees for pose 10 compared to the reference position. As seen in Fig. C.8, parts of the unaligned point clouds (left) are missing from the aligned point clouds (right). These parts are excluded either due to detected errors or simply due to occlusion and shadowing. The mean of the median values of these point errors is 0.19 mm ($\pm$SD = 0.02 mm) for the two poses shown in Fig. C.8(right). It is noted that the per point errors for the two poses are similar even though pose 10 is one of the outlying poses.

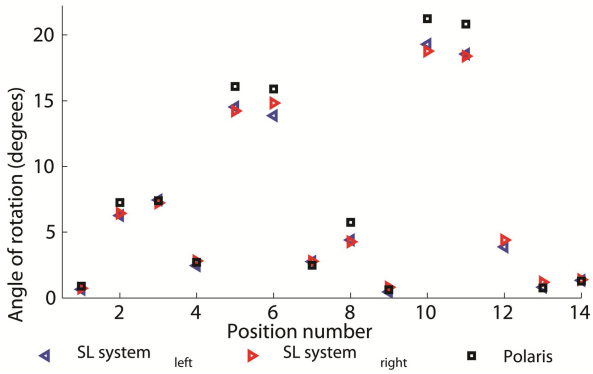


Figure C.7: Comparison of the left SL system, the right SL system, and the Polaris Vicra system on a human test subject ; the estimated relative rotation to a mean rotation for fourteen positions. No Polaris data for pose 12 due to too few markers registered by the Polaris sensor.

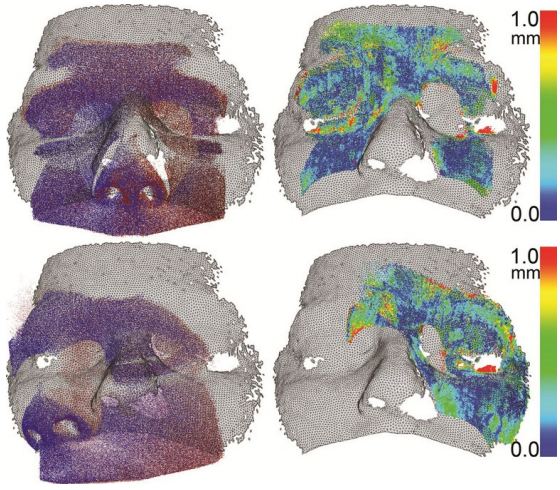


Figure C.8: Human surface scans inside the HRRT PET scanner in pose 7 and pose 10. These poses almost correspond to the positions shown in Fig. C.1. 3D point clouds before (left images) and after (right images) the ICP alignment of the left (blue) and right (red) system. The colors on the right images represent the distance [mm] to the target.

Finally, we demonstrate that the NIR illumination has the same performance for surface registration as visual light illumination. Figure C.9 shows results of the NIR-SL system for a human volunteer. This experiment was conducted outside the HRRT PET scanner with the volunteer sitting upright in a chair with support for the back and not for the head. The centroids of the reconstructed 3D point clouds from the four positions were 31–46 mm from the centroid of the reference target. The main motions were axial/downward translation and sidewise rotation which are typical motions seen in PET brain imaging. The captured images are non-blurred and have high contrast (utilizing approximately 80% of the 8 bit colors). This supports the high quality of the captured images using the NIR LED. The reconstructed 3D point clouds correspond to the regions of the scanned surface. The errors of the ICP alignment are less than 0.5 mm for the majority of regions. The mean of the median values of these point errors are 0.29 mm ($\pm$SD = 0.09 mm)(compared to 0.19 mm ($\pm$SD = 0.02 mm) for the visual light illumination). Small areas especially at the borders and the nose have errors > 1 mm seen as the red areas (Fig. C.9).

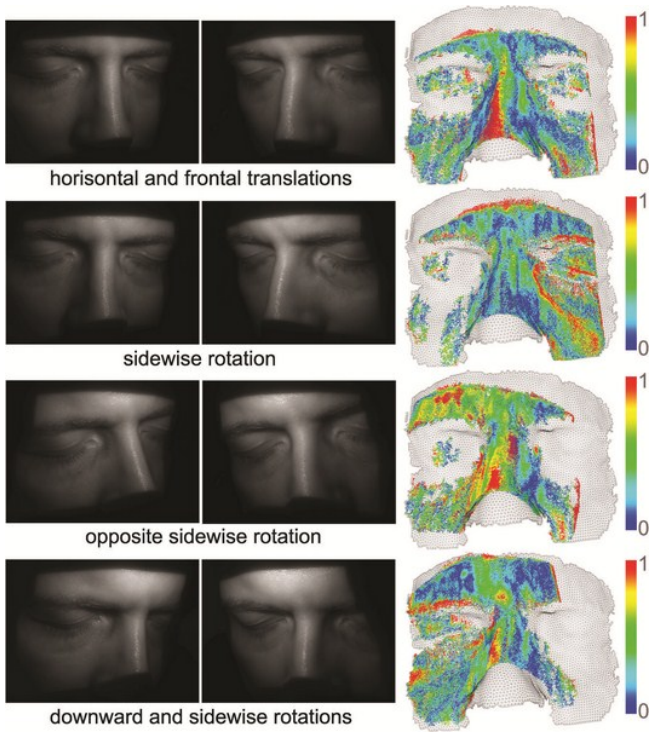


Figure C.9: NIR-SL results. Left: Captured regions with the two cameras. Right: Errors [mm] of the alignment on top of the triangulated reference target.

C.5 Discussion

A structured light system developed to register 3D head motion was applied to a clinical setting and shown to work given narrow scanner geometries such as the HRRT PET scanner. The design is miniaturized, flexible, and does not need any markers. These qualities make the system usable and valuable in a clinical setting. A markerless system is also timesaving and hygienic in terms of hospital use. Another important advance of not using markers is the elimination of the major source of error when attaching markers onto the head. When using the Polaris Vicra system there is potential risk that the tracking tool moves relative to the skull either because of the attachment or movements not related to the skull such as facial movements. Facial movements can obviously also introduce tracking errors with the SL system. However, this system relies on thousands of points thereby improving the robustness of the transformation estimation.

The structured light system was realized in both visible and NIR versions, SL and NIR-SL, respectively. The two systems were compared to the Polaris Vicra system on a phantom set up in order to get ground truth motions. The accuracy of the SL system was equal to the Polaris Vicra system with a RMS error of 0.09 degrees for axial rotation from -20 to 20 degrees and a RMS error of 0.26 mm for translation over a range of 50 mm (Fig. C.6).

This result is supported by the accuracy found in [111] for a similar experiment. Thus on a rigid object where the tool is fixed the accuracy of both systems are in the order of a few tenths of a millimeter for the brain region, e.g. if a head is rotating around the point of contact in the back head the displacement of the frontal brain lobe 10 cm away for a 0.09 degrees rotation is $\tan(0.09^\circ) \cdot 10$ cm = 0.16 mm.

We believe this accuracy is representative of the achievable accuracy on humans since some aspects can be improved such as facial movements, time of image capturing, and combining the left and right systems as explained below. The accuracy are an order of magnitude less than the current resolution of the HRRT PET scanner.

A human volunteer study was performed inside the HRRT PET scanner to demonstrate the SL and NIR-SL systems usability for the HRRT PET scanner on humans. We have shown detailed 3D point clouds of the face surfaces which are aligned to the target with a median of the per point error of around 0.2 mm (Fig. C.8). These errors are due to reflectance of the light, noise, and motion. The output transformations from the SL system are based on thousands of points and thus the accuracy of the rigid body transformation is expected to be $<<$ 1 mm. From the comparison of the systems in fourteen head poses we

find twice as good agreement between the two SL systems (SD(SL_{left}-SL_{right}) = 0.41^o) compared to the agreement between the Polaris Vicra system and one of the SL systems (SD(Pol-SL_{left}) = 0.85^o and SD(Pol-SL_{right}) = 0.97^o). Some poses have significant larger difference with the Polaris Vicra system (Fig. C.7) for which we have no unique explanation. It could be due to the calibration of the SL system or motion of the Polaris tracking tool. We observed that the per point alignment error (Fig. C.8) was not related to the size of the performed motion. The difference between the left and right SL system is at the same level as for the phantom experiment if we do not only consider the system with the best angular position relative to the surface. Therefore, the accuracy of a total SL system is improved by combining information from the left and right system taking the angular distortion into account or simply excluding point clouds with normals perpendicular to the image axis.

We have presented a miniaturized NIR-SL system for 3D head tracking. In this system, the visible LED of the original Texas Instruments Pico DLP was replaced with a NIR LED. We demonstrated the system on a human volunteer for four different head poses. The 3D point clouds are highly detailed as for the visible system. The median of the aligned per point errors is 50% larger than for the human volunteer study with the visible SL system. This could be explained by the differences between the experiments. In the NIR-SL case, the volunteer was sitting in a chair without head support, whereas for the SL case the volunteer was lying in the HRRT PET scanner with head support. This set up might have introduced some motion during the image capturing of a set of images of 3–4 seconds. The rate of the image capturing has not been optimized for these prototypes of the SL system.

Based on the detailed point clouds, the results of the aligned point clouds, and comparing the quality of the captured images we expect to obtain the same accuracy of this system compared to the visible SL system. The results with the NIR-SL system are representative for the experiments we have performed inside the HRRT PET scanner. This error indicates that the system will be able to accurately determine the pose changes for real humans in a clinical environment with simultaneous PET acquisition. In order to realize motion compensated HRRT PET imaging based on the NIR-SL system time synchronization between the HRRT PET scanner and the NIR-SL must be established, the HRRT PET and NIR-SL coordinate system must be aligned, the NIR-SL must be configured to run continuously, and an algorithm for motion compensated reconstruction must be in place. Basically, two approaches can be taken for motion compensation depending on the nature of the motions of the study subjects. The multiple acquisition frames method assumes that the subject lies still for longer periods of time interrupted by short periods of motion [123]. The motion tracking device is used to identify these periods of no motion. For each such period a PET image (frame) is reconstructed using standard algorithm.

The frames are then combined to a reconstructed PET image in a second step. An alternative method does not make an assumption of the motion pattern but requires continuous estimates of pose [22]. In this case each PET event (LOR) is repositioned before using a standard reconstruction method.

C.6 Conclusion

We present a compact vision system based on a customized Texas Instruments Pico DLP projector fitted with a NIR LED. The system is adapted for motion correction in high resolution PET brain imaging. While the system's accuracy is comparable with the current state-of-the-art optical trackers, it is more flexible and the system is fully automatic and does not rely on markers. Furthermore, the NIR LED ensures a more comfortable experience for the patients. This is a step toward a fully automatic tracking system designed for the HRRT PET brain scanner, but with potential use in other scanners and imaging modalities where an external tracking system is currently needed.

Acknowledgment

The authors would like to thank Siemens Healthcare A/S and The Danish Agency for Science, Technology, and Innovation, which funded the work. We would also like to thank Rigshospitalet for making clinical scanner facilities available. The John and Birthe Meyer Foundation is thanked for the generous donation of the HRRT PET scanner. Finally, we would like to thank the staff at The Mechanical Workshop at Rigshospitalet who helped us produce the mechanical devices.

Appendix D

3D Surface Realignment Tracking for Medical Imaging: A Phantom Study with PET Motion Correction

Oline V. Olesen, Rasmus R. Paulsen, Rasmus R. Jensen, Sune H. Keller, Merence Sibomana, Liselotte Højgaard, Bjarne Roed, and Rasmus Larsen.

Abstract

We present a complete system for motion correction in high resolution brain positron emission tomography (PET) imaging. The system is based on a compact structured light scanner mounted above the patient tunnel of the Siemens High Resolution Research Tomograph (HRRT) PET brain scanner. The structured light system is equipped with a near infrared diode and uses phase-shift interferometry (PSI) to compute 3D point clouds of the forehead of the patient. These 3D point clouds are progressively aligned to a reference surface, thereby giving the head pose changes. The estimated pose changes are used to reposition a sequence of reconstructed PET frames. To align the structured light system with the PET coordinate system, a novel registration algorithm based on the PET transmission scan and an initial surface has been developed. The performance of the complete setup has been evaluated using a custom-made phantom, based on a plastic mannequin head equipped with two positron-emitting line sources. Two experiments were performed. The first simulates rapid and short

head movements, while the second simulates slow and continuous movements. In both cases, the system was able to produce PET scans with focused PET reconstructions. The system is nearly ready for clinical testing.

D.1 Introduction

Patient head movement during high resolution brain positron emission tomography (PET) scanning will cause blurring and ghosting [58]. The low count rate and resulting low contrast makes it almost impossible to perform motion correction on raw PET data, and therefore most methods rely on external tracking of the head movement [123, 129, 132, 167]. The Polaris Vicra (Northern Digital Inc.) tracking system has been used as the reference on many PET installations [93]. While the Polaris system is well tested and accurate, it suffers from problems related to attaching optical markers to the patient's head. Experience shows that in a clinical setting, the markers are difficult to attach such that they stay in position during the entire scan. A markerless system that fits into the narrow PET tunnel will improve the clinical acceptance and the diagnostic value of PET brain scans.

We have previously described a structured light based system that is based on a small projector and two small cameras [111] for tracking patient head pose. This system has been modified as described below and it is referred to as Tracoline. The Tracoline system has been designed to fit into the patient tunnel of the Siemens High Resolution Research Tomograph (HRRT) PET brain scanner. The HRRT PET scanner has a spatial resolution down to 1.4 mm [116] and is therefore well suited for testing new motion correction methods. The Tracoline system is based on the progressive reconstruction of 3D surfaces of the upper face region of the patient in the scanner. The pose changes are found by computing the rigid transformation between the current scan and the initial surface scan. The system described in [111] was based on visible light and did not operate in real time. Visible light scanners are not suited for repeated human facial scans. Furthermore, to be functional the system needs to acquire frames sufficiently fast to faithfully capture patient head movements. In this paper, we describe a system using invisible light with a camera acquisition rate of 30 frames per second.

While the previous paper focused on measuring the accuracy of the structured light tracking system using a rotation stage as ground truth [111], the real interest is the improvement of the PET scans. To be able to evaluate the quality improvement of the PET scan, a scan using a radioactive tracer must be performed. A common approach is to use a phantom and compare the resulting

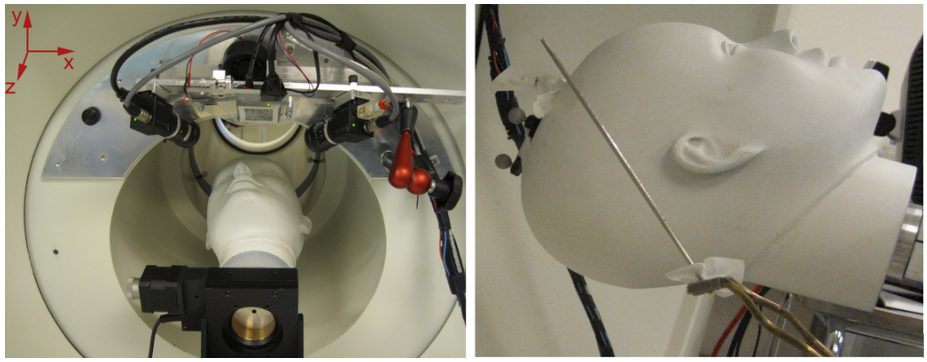


Figure D.1: Left: The patient tunnel of the HRRT PET scanner with the Tracoline system mounted. The phantom is mounted on a rotating stage rotated to the right (-10 degrees). Right: The phantom including one of the two radioactive line sources. It is placed in the head in the same angle as shown here to simulate the brain.

PET scan with the known geometry of the phantom [116]. We have therefore designed a customized phantom with a radioactive source and used this in the evaluation of the Tracoline based motion correction. Our system setup with the Tracoline system and the custom phantom can be seen in Fig. D.1. Compared to other external tracking systems, where the geometric alignment between the tracking system and the PET scanner can be problematic, we investigate a novel alignment approach based on aligning the Tracoline system scan directly to the PET transmission scan.

D.2 Experiments and Methods

The Tracoline system consists of two Point Grey Flea2 cameras (1288×964 pixels), each running at 30 frames per second. The Pico Digital Light Processing (DLP) projector from Texas Instruments is used to project phase-modulated patterns. One of the light diodes of the Pico projector has been replaced with a near infrared (NIR) diode resulting in a non-visible projected pattern. The projector is controlled by a GFM Pico developer kit board that also sends trigger pulses to the cameras, in order to synchronize the projected patterns and the shutter of the cameras. A multi-threaded C++ program running on a standard portable computer acquires the real time camera data and stores them as image files. The 3D point cloud generation, surface reconstruction, and alignment are done in a post-processing step. The Tracoline system and HRRT PET ac-

quisition computer are synchronized through an internal network time protocol (NTP) server.

The 3D point cloud generation is based on phase-shifting interferometry (PSI) [73] where a set of 2D interferograms are projected and projector-camera correspondence can be found using phase unwrapping. This is explained in detail in [73, 111]. While three patterns are used in [111], the system is now extended to use six patterns with varying wavelengths to make the phase unwrapping more robust to discontinuities in the surface. Since each point cloud computation requires six frames, the effective tracking frequency is 5 Hz. A surface is reconstructed using a modern algorithm [117] based on the point cloud acquired in the initial position. The facial pose changes are then found by rigidly aligning the following surface scans to this reference surface using an optimized, iterative, closest point (ICP) algorithm [166].

To correct for motion, we need to know the transformation between the HRRT PET scanner coordinate system and the Tracoline system. To estimate this transformation, we use the transmission scan of the HRRT PET scanner, which is also used for the attenuation and scatter correction within normal PET reconstruction. The transmission scan is a voxel volume similar to a computed tomography (CT) scan. The initial reference surface scan is captured by the Tracoline system during the transmission scan, thus creating correspondence. The transformation is computed using a pseudo-ICP algorithm [139]. The surface scan is scaled to fit the volume, and manually rotated and translated into an initial position. To find correspondence between surface sample points and the volume, the volume is sampled in the normal direction (both positive and negative) of the surface scan to find the point with maximum gradient. Knowing the general orientation of the patient in the PET scanner, we use the absolute gradient in the x (left/right) and z (axial) direction and the negative gradient for the y (anterior/posterior) direction:

$$\Delta f = \left|\frac{\partial f}{\partial x}\right| - \frac{\partial f}{\partial y} + \left|\frac{\partial f}{\partial z}\right|$$

With a point correspondence, a rigid transformation is found using the closed-form loop to estimate the absolute transformation [70]. With an initialization, transformation, this process is iterated until the transformation of the Tracoline scan converges to the volume data.

In order to apply the motion correction to the PET data, we apply the multiple acquisition frames (MAF) method [123]. In [116] the MAF method was demonstrated on the Siemens HRRT PET scanner using the tracking input from the Polaris Vicra system. We divide the PET emission list mode data into equal time length intervals and, for each interval, a PET frame is reconstructed using the 3D ordered subset expectation maximization (3D-OSEM) algorithm with

resolution modeling and incorporating a spatially invariant point spread function [154]. These frames are then repositioned to a reference position using the Tracoline tracking system.

A custom phantom with known geometry was designed. It consists of a hollow plastic mannequin head with a very low attenuation coefficient. Two radioactive line sources are inserted into the head to provide activity for the HRRT PET scanner. The activities of the line sources are 2×7.8 MBq each, created by a positron-emitting germanium-68/gallium-68 generator. As can be seen in Fig. D.1, the line sources go through the head from the back of the skull to the forehead. The phantom was mounted onto a rotation stage by Thorlabs and placed in the patient tunnel of the HRRT PET scanner.

Two experiments where performed using the stage to rotate the head. In experiment one, the head was rotated from -20 to 20 degrees in steps of 5 degrees. At each position a 30 s frame was PET reconstructed and repositioned. Data with motion was excluded from the reconstruction. In experiment two, the head was rotated from -10 to 10 degrees in a continuous motion with a maximum speed of one degree per second. The PET data was reconstructed using one second frames in experiment two. Experiment one simulates the clinical situation where the patient is performing a rapid head motion followed by a stationary period. State-of-the-art practise is to discard PET data during such rapid motions. The second experiment simulates e.g. a patient falling asleep, where the head is slowly drifting from side to side.

We evaluate the effect of the motion correction on the reconstructed PET images by calculating Dice's coefficient (percent volume overlap) [33] between a reference image recorded without phantom motion, the motion distorted image, and the Tracoline based motion-corrected image. The number of voxels, N, included in the calculation is set to a value corresponding to the number of voxels inside the tubes 2.5 times the diameter of the PET sources used (outer diameter 3.2 mm and active length 168 mm) [116]. The extended volume is used in order to compensate for partial volume effects. In each image to be studied, the set of the N most intense voxels is extracted and used for the Dice's coefficient computation, presented as the percentage of overlapping voxels. In addition we compute the normalized cross correlation between the reference image and each image frame, either motion-corrected or uncorrected [56].

D.3 Results and Discussion

The rigid transformation between the coordinate system of the Tracoline system and the PET image frame is obtained from using the described surface-to-volume alignment. Figure D.2(a) shows the reference surface aligned to maximum gradient points in the transmission scan.

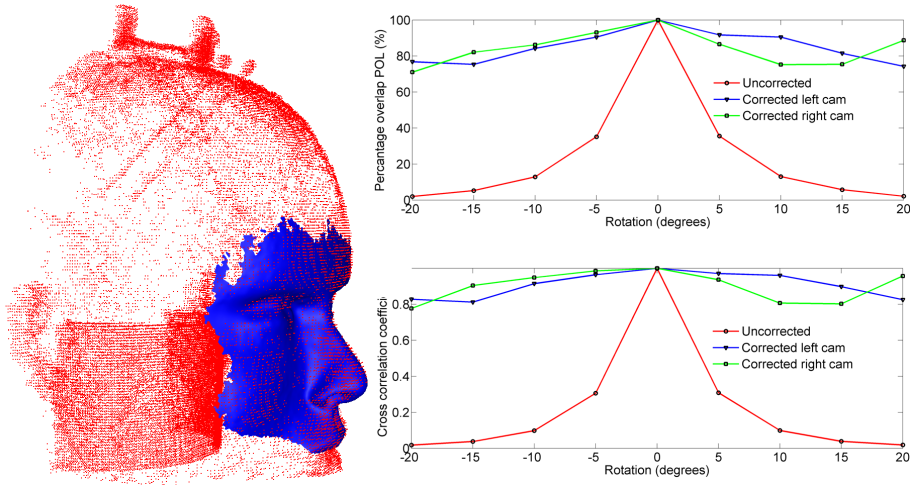


Figure D.2: Figure D.2(a) shows the alignment between maximum gradient points in the transmission scan shown as red dots and the Tracoline face scan shown as a blue surface. Figure D.2(b) shows quantitative results of the stepwise experiment. Top: the percentage of overlapping points in the reference PET image compared to the unaligned/aligned PET images based on either the right or the left camera. Bottom: the cross correlation between the reference PET image and the unaligned/aligned PET images.

Figure D.2(b) shows the results of the first experiment with stepwise rotation of the phantom. The top figure shows the percentage overlap between PET frames of the line sources in the reference position and a scan position as a function of the performed rotation of the head. Similarly, the bottom plot of Fig. D.2(b) shows the correlation coefficient between the reference image and a motion-corrected/uncorrected image for the different scan positions. Results based on tracking information from the left and right camera of the Tracoline system are shown in green and blue colors respectively, while the red curve represents the uncorrected image results. The overlap and the correlation measures are in agreement. The results of the uncorrected frames decrease with the size of performed rotation from an overlap of 100% down to 2% at ± 20 degrees. The overlap of the motion-corrected reconstruction is improved significantly

for all positions with percentage overlap of 71–93%. The overlap is not 100%, which is mainly due to the internal calibration of the Tracoline system, the ICP alignment, and the geometrical alignment with the HRRT PET scanner. In addition, the interpolation error, combined with the straight and narrow line sources of the phantom (with a diameter similar to the voxel size of 1.2 mm), induces partial volume effects and thereby decreases the overlap and correlation measure. The differences between the left camera and the right camera could be explained by the construction of the reference surface scan, where left camera was chosen as the basis. The result is similar to [116], where the overlap was 65–85% for a 10 degrees corrected rotation. However, the two studies cannot be directly compared since the phantom designs are different.

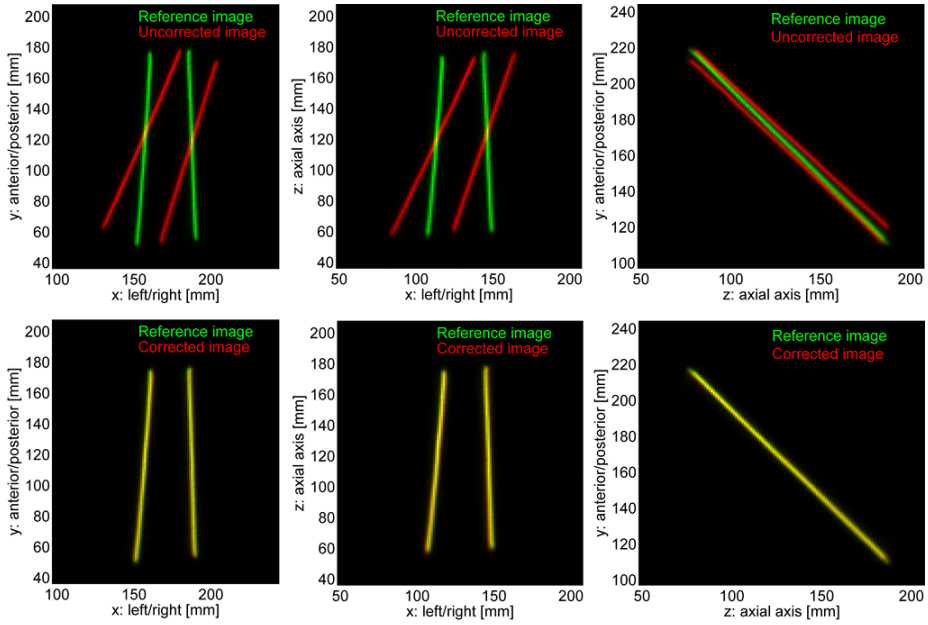


Figure D.3: The figure shows the summation of the PET images along 3 different axes for a reference image (shown in green) and a target image (shown in red) with a 20 degree rotation (overlap = 89%, shown in yellow). The uncorrected image is shown as captured in the first row, while the second row shows the image after motion correction.

A visual evaluation of the motion correction is shown in Fig. D.3 for the maximum rotation of 20 degrees. The PET images are summed along one dimension and visualized on top of each other pairwise in the red and green color channels. The overlapping pixels of the two PET images appear yellow. The top row of the figure shows the reference image and the uncorrected image as two sets of rods rotated approximately 2 cm at the end points. These correspond to the

relevant brain regions: the frontal lobe and cerebellum. The bottom row shows the reference image and the corrected image seen as two yellow rods, demonstrating a near-perfect motion correction. This position has an overlap of 89% in contrast to the rotation of -20 degrees with an overlap of 71%.

The results of the second experiment, with a continuous rotation of 20 degrees demonstrating the realtime pose registration of the Tracoline system, are presented in Fig. D.4. The one second PET frames are summed and fused with the transmission image of the phantom. The top of the figure shows a row of uncorrected images, where the motion of the line sources is seen as blurred circle parts. The bottom row shows the motion-corrected image, where the previously blurred parts appear in focus and with high intensity. The cross section of the line sources shows dots with a diameter of only a few pixels. Long drift motion is a very complex problem to overcome using image registration methods for motion estimation, and this is why an external tracking system is of great value. Our latest results show that continuous motion can be tracked in real time and PET frames successfully corrected.

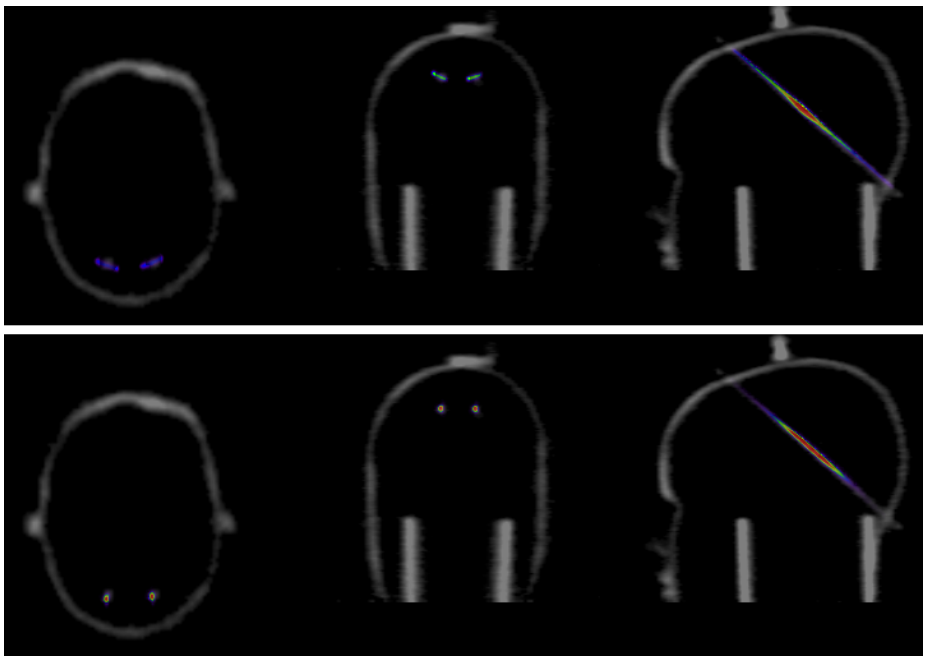


Figure D.4: Results of the dynamic PET scan. One-hundred, one-second frames uncorrected (top) and MAF motion-corrected (bottom) are summed and fused with a transmission scan. The frame repositioning is based on the left camera alignment.

D.4 Summary and Conclusions

This paper describes a complete system for motion correction in high resolution PET brain imaging. It is based on a small and flexible structured light scanner mounted above the patient tunnel of the PET scanner. The scanner is equipped with a near infrared light source, making it suitable for future patient examinations. Furthermore, the system tracks the head pose changes with a frequency of 5 Hz, which is suitable for the head movement experienced during real clinical PET scanning. In order to align the two systems, a novel algorithm using the HRRT PET transmission scan and the initial surface scan was presented. The performance of the system was evaluated using a custom-designed phantom with two radioactive line sources mounted on a programmable rotation stage. The results of the two experiments are very promising. The first experiment simulates rapid but short head movements and the second experiment simulates slow but longer head movements. Quantitative analysis shows that the combined system is able to robustly reduce motion artifacts and greatly improve PET scans for scenarios involving both slow and rapid movements. The system is nearly ready for actual clinical testing.

Acknowledgements The John and Birthe Meyer Foundation is thanked for the generous donation of the HRRT PET scanner and the cyclotrons for isotope production.

Appendix E

List-mode PET Motion Correction Using Markerless Head Tracking: Proof-of-Concept with Scans of Human Subject

Oline V. Olesen, Jenna M. Sullivan, Tim Mulnix, Rasmus R. Paulsen, Liselotte Højgaard, Bjarne Roed, Richard E. Carson, Evan D. Morris, and Rasmus Larsen.

Abstract

Purpose: Precise head motion registration is crucial for correct motion correction (MC) in positron emission tomography (PET) brain imaging. State-of-the-art tracking systems applied with PET brain imaging rely on markers attached to the patient's head. The marker attachment is the main weakness of these systems. Attachment failures cause loss of motion information and require dependence on software methods, which must be validated for each tracer. The purpose of this study is to demonstrate a previously developed custom designed markerless tracking system for improved motion correction on PET brain images.

Method: A healthy volunteer participating in a cigarette smoking study to image dopamine release was scanned twice for two hours with ^{11}C-racolopride on the high resolution research tomograph (HRRT) dedicated brain scanner. Head motion was independently measured, with 1) a commercial marker-based device (Polaris Vicra, NDI), and 2) the proposed stereo vision-based system (Tracoline). A list-mode event-by-event reconstruction algorithm using the detected motion was applied. To investigate the tracking performance, the data were evaluated based on similarity between frames using cross-correlation (XC), similarity of time-activity-curves (TACs), and a two-class discrimination approach quantified with the Mahalanobis distance (MD) measure. The tracking outputs were directly compared with transformations of a test-point in the left striatal region. A phantom study with similar procedures and with hand-controlled continuous random motion was obtained.

Results: Motion of the left striatal region was time-varying in both scans with long drift motions of up to 18 mm and regular step-wise motion of 1-6 mm. Maximal tracked motions was 20–25 mm. Evaluating a test-point in the left putamen, the standard deviations (SDs) of the 3D tracking differences between the Tracoline and Polaris systems were 0.89 mm, 1.04 mm, and 0.90 mm for the two human scans and the phantom scan respectively. All the evaluated measures XCs, TACs of the putamen region, and MDs were significantly better for motion-corrected images compared to no motion correction (no-MC). The difference between no-MC and MC for the evaluated measures correlated with the tracked motion. XC and TACs from the Tracoline-based motion correction (TCL MC) and the Polaris-based motion correction (POL MC) agreed overall during the two scans.

Conclusions: A markerless tracking system was demonstrated to be applicable for PET brain imaging. The system agreed with a commercial integrated system. The patient motion during the scans was shown to affect the TACs and therefore must be corrected. Precise head motion estimation with list-mode motion correction was shown to be important for ROI analysis of small brain regions. Motion-corrected images were improved in contrast recovery of small structures.

E.1 Introduction

Improved scanner technology and reconstruction algorithms have advanced PET imaging significantly in recent decades. Time-of-flight and point spread modeling of individual PET events have improved the spatial resolution of scanners [26, 154], allowing identification of gray and white matter in the brain. However, subject motion during the PET acquisition blurs PET images and

undermines advances in high resolution scanning.

Motion correction (MC) algorithms for PET imaging date back to the 1980s [29, 100, 120, 123, 168]. Picard and Thompson proposed the Multiple Acquisition Frames (MAF) method in 1997 in which frame timing was determined based on motion information [123]. The MAF method has demonstrated its abilities well in phantoms [45,116]. More advanced methods with line-of-response (LOR) correction [14, 16, 22, 100] or correction of the system matrix have also been suggested [130]. All of these methods assume that the patient's motion during the scan is known.

Image-based registration methods using the mutual information (MI) criterion [97,153] and automated image registration (AIR) [168,169] are widely used for registration between image modalities, and have also been used for motion correction between PET frames [27, 109]. These methods do not require external measurement of patient motion. However, validation of software motion correction is required for each unique tracer distribution.

The patient's motion can be estimated in two ways; from the PET events themselves, referred to as "data-driven" motion tracking (MT), or by external MT devices. External MT systems can correct for many types of motion, such as intra-frame motion and long drift motion. It is also possible to measure motion during the transmission (TX) scan. A second TX scan can be acquired if motion is seen to corrupt the original TX. In contrast to data-driven MT, external MT estimation is unaffected by the radiotracer distribution or the image noise level.

Data-driven MT operates on time-frames of data. Within a time-frame, it is assumed that the radiotracer concentration is constant and the patient is motionless. Time-frames typically vary in length depending on the time from radiotracer injection and expected kinetics of the tracer. The time-frames can easily be 10 min in duration late in the PET acquisition in order to get sufficient image-statistics.

Further, dynamic 4D PET reconstruction techniques do not reconstruct the emission (EM) time-frames independently but utilize information from multiple EM frames to model the time-varying radiotracer distribution [124, 149]. Here, patient motion can affect the parametric model in a manner that could be confused with changes in the radiotracer distribution.

Uncorrected motion can be interpreted by the kinetic model as change in the tracer distribution thereby leading to wrongly estimated physiological parameters of interest such as tissue perfusion, blood volume, metabolic rate, receptor density, receptor occupancy, or gene expression rate [65]. 4D PET reconstruction requires external MT to separate the patient's motion from variation of the

radiotracer distribution [81, 133, 162].

A few external MT systems have been demonstrated for PET brain imaging [55, 93, 122]. A commercial Polaris tracking system from Northern Digital Inc. (NDI, Waterloo, Canada) has been favored for human PET brain imaging [93]. The Polaris system registers a rigid tracking tool of 3–6 spherical reflective markers. Alternative MT system that registers a checkerboard pattern from camera visions has been used in animal studies [174].

The drawback of existing external MT systems is that they rely on markers attached to the subject for tracking. Unfortunately, all markers to the head (e.g, helmets, goggles, band-aids, or bite fixtures) can move relative to the head [83]. Improper motion correction introduces image degeneration instead of image improvement.

There is need for a markerless external tracking system that can be incorporated into methods that reposition LORs to the patient's position at the time of the detected events. Moreover, a markerless system could improve the workflow of the PET image acquisition by the eliminating marker attachment.

We have recently presented a markerless MT system for 3D head tracking [112]. This system (named Tracoline) is based on a structured light approach to create face surface scans that are then aligned to a reference surface. The accuracy of the system was shown to be similar to the Polaris Vicra tracking system [112]. Here we demonstrate use of the markerless Tracoline system for MC in human PET brain imaging. To evaluate the system performance, a state-of-the-art list-mode event-by-event MC reconstruction algorithm [22] was applied to high resolution PET acquisition [165]. In the present "proof-of-concept" study in human and phantom scans, we show that Tracoline is able to perform dynamic head position estimation for motion correction in PET brain imaging. All scans were conducted with simultaneous Tracoline tracking and Polaris tracking and the resulting MC PET images were compared to PET images without MC.

E.2 Material and Methods

E.2.1 Human PET

Data from a healthy female smoker (50 years old), recruited for an ^{11}C-raclopride study at the Yale University PET Center to examine the effect of smoking on endogenous dopamine release, are included in this study. The study protocol

was approved by the Yale University School of Medicine Human Investigation Committee in accordance with the principles set forth in the Belmont Report. Written informed consent was obtained from the subject before the scan after complete explanation of the study procedures.

The subject was scanned twice (smoking and rest) on separate days. Each two-hour scan was completed on the high resolution research tomograph (HRRT, Siemens, Knoxville, TN, USA) [165] following a bolus plus constant infusion of ^{11}C-raclopride (Kbol = 105 minutes). During the smoking scan, the subject was asked to smoke two cigarettes in succession while supine in the scanner. Smoking commenced, 45 minutes after tracer injection. A freestanding air exhaust and filtration system was used to capture and filter cigarette smoke. A TX scan for attenuation was acquired prior to each dynamic scan for attenuation correction. Prior to PET imaging, the subject received a 3T structural MRI which was used to define regions-of-interest (ROI).

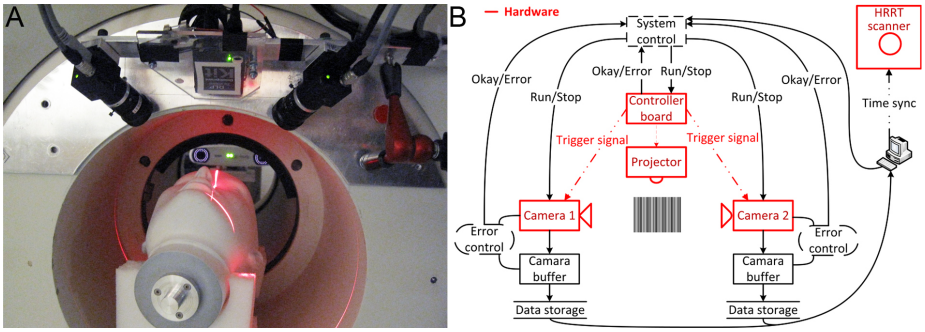


Figure E.1: A: The gantry of the HRRT PET scanner with a mannequin phantom head inside. The Tracoline system is mounted in the front above the patient tunnel and the Polaris system is seen through the gantry at the rear of the HRRT scanner. Beam of light on the mannequin face is from HRRT scanner positioning laser. B: Diagram of the Tracoline system integrated with the HRRT scanner separated into the hardware components (red) and software communications (black).

Head motion was recorded during the scan by both the Tracoline system and the Polaris Vicra system, as described in Section E.2.3. EM data were collected in list-mode. Dynamic scan data were reconstructed by the MOLAR algorithm [22] including motion correction using the information collected by either Polaris or Tracoline systems, separately. Data were also reconstructed by MOLAR without motion input. Thus, three reconstructions were studied: Tracoline-based motion correction (TCL MC), Polaris based-motion correction (POL MC), and reconstruction without motion correction (no-MC). Data were binned into three different framing schemes during reconstruction: 1×120 min,

12×10 min, and 240×30 s. These framings were used to evaluate image quality to demonstrate the effectiveness of the different motions tracking schemes within list-mode reconstruction.

Figure E.1A shows the HRRT PET scanner equipped with the two tracking systems. Both systems work with light of the same spectrum, so a visor was worn by the subject to prevent light from the Polaris system (projected from behind the scanner) from interfering with the Tracoline system.

E.2.2 Phantom PET

A phantom study was conducted to demonstrate the tracking performance of the Tracoline system on PET images with high count statistics. The phantom consisted of a hollow mannequin head equipped with two ^{68}Ge-line sources of 17.8 MBq and 2.2 MBq, respectively. The line sources penetrated the mannequin head and were fixed at the forehead and the back of the head to represent a large part of the brain (from the frontal lobes to the cerebellum). The Polaris tracking tool was fixed to the top of the mannequin head.

A five minute PET acquisition of the phantom was obtained with simultaneous Tracoline tracking and Polaris tracking (Fig. E.1A). The phantom was manually rotated and translated across a range of directions during the scan. The amplitudes of the translations and rotations were on the order of ±3 cm and ±15 degrees around the center of the head to represent motion consistent with human motion during scanning. The phantom was stationary for the first $1\frac{1}{2}$ minutes of the acquisition. The first minute of data was used as a reference period.

The PET data were reconstructed in frames of 1 min using the MOLAR reconstruction with three different motion inputs as was done with the human subject: TCL MC, POL MC, and no-MC.

E.2.3 Motion tracking

Motion during PET scans was registered with two systems; the recently developed 3D head tracking system Tracoline (TCL) [112] and the commercial Polaris Vicra (POL) system from Northern Digital Inc. (NDI, Waterloo, Canada).

The Polaris system is an optical tracking system that registers a rigid tool with 4 reflective markers at a rate of 20 Hz. The tool is mounted on the top of a

swim cap using velcro. The swim cap is then placed on the subject. In these studies, to further fix the tool in place, Coban^{TM} (a self-adherent elastic wrap) was wrapped around the head, cap, and tool without obscuring the markers. The functionality of the Polaris system is based on the assumption that the tool mounting provides a stiff coupling between the tool and the brain.

The Tracoline system is a markerless tracking system based on stereo vision. It registers motion between surface frames recovered through structured light scanning. The tracking accuracy depends on the presence of significant surface curvatures. For human head tracking, the face region around the bridge of the nose exhibits high curvatures in all directions and facial movements in this region are limited. Therefore, the middle face region is ideal for head tracking in PET brain imaging. The Tracoline system is mounted onto the front of the HRRT gantry just outside the PET field-of-view (FOV) via one or two flexible arms (Fig. E.1A).

Tracoline consists of two CCD cameras (Flea2, Point Grey Research Inc, Richmond, Canada.) and one Pico projector (Texas Instruments, Dallas, USA) modified to operate in the infrared domain to avoid subject discomfort [112]. 3D face surfaces are archived from stereo vision of one camera and the projector. Correspondence between the two image planes (projector and camera) is found by fringe projection of cosine-patterns. This method requires capture by the cameras of three patterns projected onto the subject's face. However, to achieve robust correspondence we use three cosine-patterns at two different spatial frequencies. The resulting tracking frequency is 5 Hz with an image capture rate of 30 frames-per-second for each camera. The projector is controlled through a controller board (GFMesstechnik, Teltow, Germany) which sends trigger impulses to the cameras as seen in Fig. E.1B. The controller board is connected to a PC where the raw image data are stored and the tracking estimate is obtained from post-processing. Partial face surfaces are created for each tracking frame (set of six captured images) and aligned to a reference surface using a fast iterative closest point (ICP) algorithm [166]. We select the reference surface to minimize facial movements (e.g. the eyes and mouth are excluded in order to estimate the overall 3D head motion and limit noise from facial movements). A tracking estimate is not returned if the root-mean-square (RMS) distance between the reference surface and a surface frame is > 0.6 mm.

Time synchronization: The time synchronization between the HRRT scanner and the two tracking systems was done via a network time protocol with accuracy estimated to be within half a second. This accuracy is generally adequate for human scans with poor image statistics during half a second and without regularly occurring fast motions. However, for the phantom study with constant motion faster than the accuracy of the time synchronization, the timing was more important. To overcome this, tracking inputs with different time-shifts (in

steps of 50 ms) were tested with the motion-corrected reconstruction. Tracking inputs resulting in the least blurred PET images were retained.

Motion tracking comparison: The motion estimations were compared based on a test-point in the putamen region transformed into the HRRT image frame. The 3D Euclidian distance "travelled" by the test-point was calculated relative to the mean during the TX position (or to scan start for the phantom study) [102]. The tracked positions of the test-point were resampled to achieve matching sample-points for the two tracking systems. The absolute 3D distances of the Tracoline system were subtracted from the Polaris system for comparison.

Evaluation of a transformed test-point represents both rotation and translation. A travelled 3D distance can readily be related to the movement of a brain region while inspection of the translation and rotation separately is less transparent.

E.2.4 Geometric calibration

PET motion correction must be completed in the coordinate system of the HRRT scanner. The motion is recorded in the coordinate systems of the Tracoline and Polaris, respectively. Therefore, the geometrical transformations between the HRRT scanner and the tracker systems must be determined.

The Polaris system is implemented and routinely used with the HRRT scanner at the Yale PET Center. The Polaris sensor is placed behind the HRRT scanner and registers a reference tool permanently fixed to the scanner in addition to the tool atop the subject's head. Use of a reference tool removes the restriction that the Polaris tracker is spatially fixed relative to the HRRT scanner; patient head tool tracking data are computed relative to the reference tool via appropriate quaternion arithmetic. Periodically, a Polaris system calibration is performed in order to determine the rigid body transformation between the HRRT image coordinate system and the Polaris reference tool coordinate system. This calibration procedure utilizes a simple radioactive line source phantom with rigidly attached Polaris head tool. Multiple EM scans with varied phantom location and orientation allow a fitting procedure to determine the reference tool-to-HRRT transformation matrix. Knowledge of this matrix allows the patient head tool Polaris tracking transformations to be expressed relative to the HRRT image coordinate system.

The Tracoline system is mobile and not permanently fixed to the HRRT scanner. The system is assumed only to be fixed during the scanning period. The TX scan of the subject is used for the geometrical calibration with the assumption that the subject was not moving significantly during the 6 min TX scan. The

accuracy of the geometrical calibration is much less important for the PET motion correction compared to the tracking accuracy [116]. The tracking output of Tracoline is the relative transformation to a reconstructed reference surface of the subject. This reference surface is also used for the geometrical calibration with the HRRT scanner.

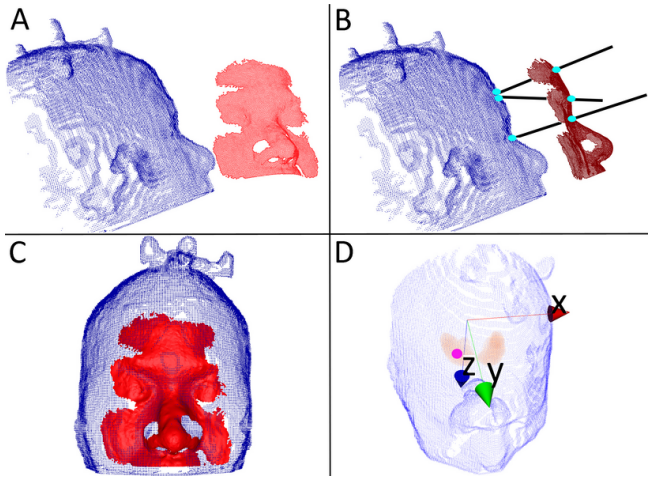


Figure E.2: TX surface points (blue) and Tracoline reference surface (red) for the smoking scan. (A) Separated view of the surfaces. (B) Principle of the point correspondence procedure; black lines represent projection along point normals and the light blue dots are the corresponding points. (C) Result of the final alignment. (D) Illustration of the test-point (pink) location used to compare Tracoline and Polaris trackings. The test-point is seen in the left putamen region. The mask used for the Mahalanobis distance measure is seen as the two red striatal regions. X, Y and Z represent the coordinate system of the HRRT scanner.

Figure E.2 shows human data used for the geometrical calibration: (blue) surface point from the TX image, and (red) corresponding Tracoline surface recorded during the TX scan. The affine transformation between the Tracoline surface and the TX surface points represents the geometric calibration denoted $\mathbf{A}_{\text{align}}$. Iterative closest point (ICP) [13, 139] where closest paired assignment is substituted by projection along normals for robustness and efficiency were used to obtain the calibration. The procedure was as follows; the TX surface points are extracted from the gradients of the TX image. Then the Tracoline reference surface is initially transformed into the TX surface followed by an iterative fitting procedure; (a) Point correspondence is found by searching for the closest TX point along the point normals of the Tracoline surface (Fig. E.2B), (b) the closed loop transformation between the two corresponding point sets is esti-

mated [70], (c) the Tracoline surface is transformed into a new position with the newly obtained transformation and the process is repeated until convergence (Fig. E.2C).

E.2.5 Motion correction

The MOLAR reconstruction [22] is used as a basis for this study. The algorithm is a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction developed for the HRRT scanner. MOLAR assumes that motion is known and the end points of each LOR are repositioned based on the motion input. The motion input is a time-series of rigid transformations that describe the transformation back to a reference position (the average position during the transmission scan) in the HRRT scanner frame at every time point. A rigid transformation in the Euclidian space is represented by a 4×4 matrix that consists of a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{E.1}$$

The MOLAR motion input is multiple such linear sub-transformations considering the brain as a rigid mass fixed to the skull. Let $[\ x_i \quad y_i \quad z_i \quad 1\ ]^{\mathrm{T}}$ be an image point in homogenous coordinates in the HRRT coordinate system and $[\ x_n \quad y_n \quad z_n \quad 1\ ]^{\mathrm{T}}$ be the same point in the tracker coordinate system. The points are transformed from the tracker coordinates to the HRRT coordinates with the transformation from Section E.2.4:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = \mathbf{A}_{\mathrm{align}} \begin{bmatrix} x_n \\ y_n \\ z_n \\ 1 \end{bmatrix} \tag{E.2}$$

The motion is registered in the tracker coordinate system and the output of the tracking system is the transformation $\mathbf{A}(t)$ that transforms the origin of the tracker coordinate system into the tracked position (e.g. Polaris tracking tool). The registered motion between a reference position $[\ x_n^*(t) \quad y_n^*(t) \quad z_n^*(t) \quad 1\ ]^{\mathrm{T}}$ and a position at a given time in the tracker coordinate system $[\ x_n(t) \quad y_n(t) \quad z_n(t) \quad 1\ ]^{\mathrm{T}}$ is then:

$$\begin{bmatrix} x_n^*(t) \\ y_n^*(t) \\ z_n^*(t) \\ 1 \end{bmatrix} = \mathbf{A}_{\mathrm{ref}} \mathbf{A}(t)^{-1} \begin{bmatrix} x_n(t) \\ y_n(t) \\ z_n(t) \\ 1 \end{bmatrix} \tag{E.3}$$

where $\mathbf{A}_{\mathrm{ref}}$ and $\mathbf{A}(t)$ are the output of the tracking system and the subscript ref refers to the transformation at the reference position. The motion of a point

i to its reference position in the HRRT coordinate system at a given time can be derived from (E.2) and (E.3):

$$\begin{bmatrix} x_i^*(t) \\ y_i^*(t) \\ z_i^*(t) \\ 1 \end{bmatrix} = \mathbf{A}_{\text{align}}\mathbf{A}_{\text{ref}}\mathbf{A}(t)^{-1}\mathbf{A}_{\text{align}}^{-1} \begin{bmatrix} x_i(t) \\ y_i(t) \\ z_i(t) \\ 1 \end{bmatrix} \tag{E.4}$$

The MOLAR motion input is the product of the four sub-transformations in (E.4). The Tracoline system estimates the transformation between a reference surface and a surface frame. Thus choosing the position of the reference surface as the reference position for the PET motion correction gives $\mathbf{A}_{\text{ref}}\mathbf{A}(t)^{-1}$ directly from the ICP alignment.

In the MOLAR reconstruction [22], the end points of each LOR are repositioned based on the current motion information. Since data are mapped back to the transmission scan position, a fixed attenuation map is used. However, normalization and randoms estimates are based on the actual detectors involved in each coincidence. Further, since the LOR sampling, i.e., the system matrix, differs from frame to frame, a unique algorithm normalization term is calculated for each frame based on the motion data.

E.2.6 Cross-correlation

In order to quantify the effects of motion correction on the reconstructed images, a similarity measure to a reference image frame was estimated. The cross-correlation (XC) of reconstructed 30 s image frames in 3D to the image frame from a reference period was calculated [56] within a brain mask. The brain mask was generated by thresholding, hole-filling, and dilation of a PET frame for the entire two-hour sequences with no-MC.

For each of no-MC, TCL MC, and POL MC reconstruction the correlation reference is the mean of the set of 30 s frames of the individual reconstructions during the same period with limited motion. The reference period was chosen based on the tracking estimate and a visual inspection of the raw image stream captured by the Tracoline system. The XC measure quantifies the similarity between frames including the changes from the tracer-variation.

E.2.7 Mahalanobis distance

The quality of the reconstructed PET images as a function of motion correction procedure was quantified by a second measure. We used a simple model of two populations to show the improvement in image quality due to motion correction. The populations were tissue with high radiotracer uptake and tissue with low or no radiotracer uptake. We focused only on a brain region with high uptake during a prolonged period of time (10–60 min); the caudate nucleus and putamen, including surrounding areas (at early times, there are other regions with high uptake). The PET images were masked to a ROI where the volumes of the high uptake and low uptake regions were approximately equal. We modeled the PET response as being of uniform intensity inside each region. The emission events of PET imaging are in their nature Poisson distributed. Due to the reconstruction process of PET images the response is typically a right-skewed distribution with a significant over-dispersion relative to the Poisson distribution. In order to compensate for the right-skewed distribution a square root transformation is applied. For true Poisson data this would to first order transform the data to a Gaussian distribution. We use a simple two population mixture-of-Gaussian model to examine the effect of motion correction. If motion is present then we expect that blurring occurs in the spatial interface between the two populations. This will pull the mean values of the Mixture-of-Gaussians fit to the overall mean. Therefore the Mahalanobis distance (MD) between the means of the two fitted Gaussians will be smaller for the no-MC case under motion than for the motion-corrected images.

The Mahalanobis distance gives high values for well separated means with small variances. The model was fitted to the histogram using the maximum-likelihood expectation maximization algorithm [32]. The MD measure is calculated as

$$MD = \frac{(\hat{\mu}_1 - \hat{\mu}_2)^2}{\frac{1}{2}(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)}, \tag{E.5}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ and are the estimated mean and estimated variance for a given two normal distribution.

E.2.8 ROI analysis

The XC and Mahalanobis analyses are data driven and do not make strong assumption on biological variation in specific compartments. However, the prior acquisition of structural MRI allows for inspection of the image reconstruction of segmented anatomical regions. Caudate, putamen, and cerebellum ROIs were taken from the AAL (Anatomical Automatic Labeling) template. Dorsal

caudate, dorsal putamen, and ventral striatum ROIs were defined based on guidelines from Mawlawi et al. [98]. Summed PET images from 0–10 minutes were registered to the subject's T1-weight MR image, which was registered to an MR template using a 12-parameter affine transformation. TACs were extracted from the 240×30 s reconstructions. TACs were filtered with a local least squares regression (loess) method [25] to smooth the dense sampled TACs.

E.3 Results

In the human subject scans, injected activity and mass were 654 MBq and 1.52 μg, respectively, for the smoking scan and 758 MBq and 3.02 μg for the rest scan. Figure E.3 shows PET images from the three reconstructions of the two-hour smoking scan.

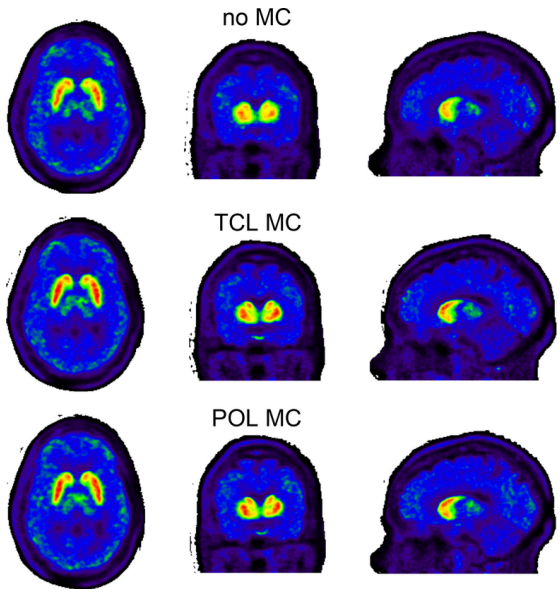


Figure E.3: PET images of two-hour acquisition (smoking scan); left-to-right axial, coronal, and sagittal images. Top: reconstruction without motion correction. Middle: Tracoline-based motion correction. Bottom: Polaris-based motion correction. The color scale is the same on all images. Compared to the motion-corrected frames, the top row with no-MC exhibits blurring and lack of detail in the high uptake region around the caudate nucleus and putamen.

There is a high uptake in the caudate nucleus and the putamen, as would be expected with ^{11}C-raclopride (a D2/D3 receptor antagonist). The half-moon shape of the caudate and putamen is most easily visible on the sagittal slices of the MC images. Separation of the caudate head and putamen is noticeable on the coronal MC images but less so on the no-MC images due to blurring. There is better contrast in the motion-corrected images and less blurring than in the non-corrected images.

Motion tracking comparison. Figure E.4 shows the motion of the test-point in the HRRT coordinate system tracked with the Tracoline system (green) and the Polaris system (red), respectively. The two test-points, one for each scan, are within the left striatum region as seen in Fig. E.2D for the smoking scan.

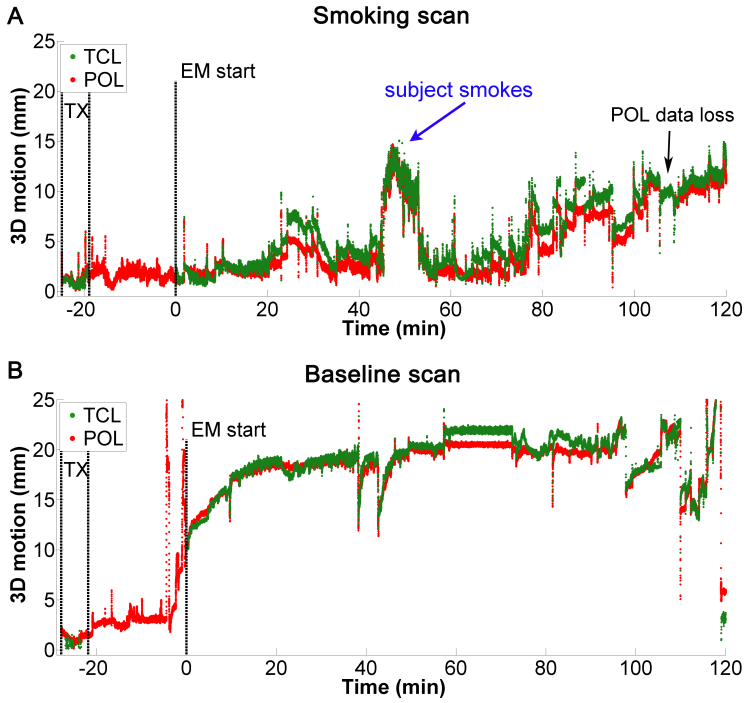


Figure E.4: Top: Tracked 3D motion of the test-point within the left striatum as seen in Fig. E.2D during the PET acquisition. Green and red represent Tracoline and Polaris trackings, respectively. The tracking information is shown in the coordinate system of the HRRT scanner. The mean and SD of the tracked motion (as plotted) during the EM acquisition were; μ_{TCL} = 6.03 mm, μ_{POL} = 4.91 mm, SD_{TCL} = 3.51 mm, and SD_{POL} = 3.35 mm for the smoking scan and μ_{TCL} = 19.3 mm, μ_{POL} = 18.7 mm, SD_{TCL} = 2.60 mm, and SD_{POL} = 2.80 mm for the baseline scan.

The test-point moves 10–15 mm during the smoking period (45 min after time of injection). There is also significant motion (> 4 mm) unrelated to smoking at 21–31 min and after 77 min (Fig. E.4A). The motion during the baseline scan is mainly drift (of 15–18 mm) which occurs at the beginning (first 15 min) of the EM scan (Fig. E.4B). Besides this, there are only a few peak motions (4–6 mm) which occur at 38 min and 43 min after time of injection. There is only limited motion during the EM scan while there is a large motion of 1.5–2.5 cm between the TX and EM acquisitions. The standard deviations (SDs) of the 3D tracking differences of Tracoline and Polaris are 0.89 mm and 1.04 mm for the smoking and the baseline scans, respectively. The Polaris tracking tool failed to register data at 105–110 min. The Polaris system reported 'too few markers' indicating that one or more markers were obscured from the Polaris sensor during this time. This could be due to slipping of either the visor or the cap.

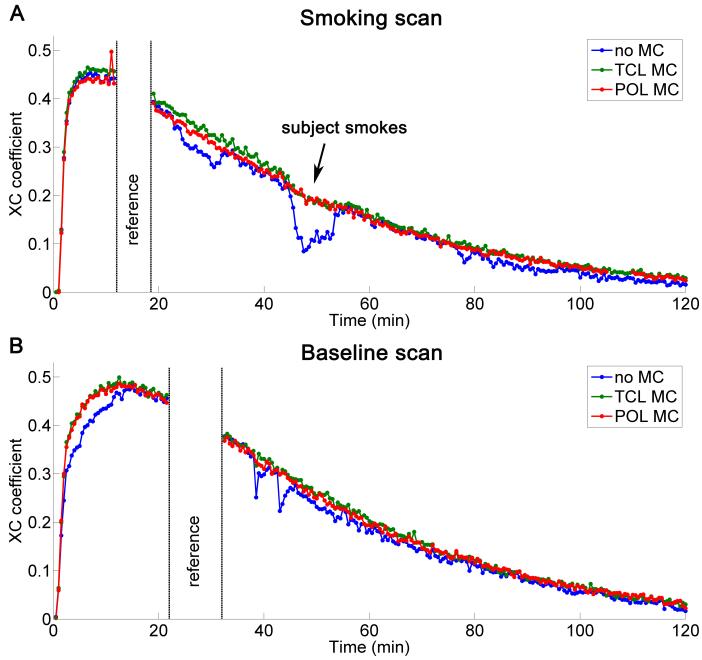


Figure E.5: Cross-correlation(XC) measure of the 30 s frame images for the smoking (top) and baseline scans (bottom). The reconstructed frames are correlated with the mean of a sequence of frames from the scan itself referred to as the reference period. The reference periods are frames 29–37 (14–18.5 min) and frames 44–64 (21.5–32 min) for the smoking and baseline scans, respectively. These periods were chosen as references because tracked motion was low. At times with large motion (cf. Fig. E.4) the reconstructed images with no-MC (blue) exhibit smaller correlation. The general shape of the curve reflects the TACs in Fig. E.6.

Cross-correlation. The XC as a function of time is shown in Fig. E.5 for the smoking scan (top) and the baseline scan (bottom). The overall shape of the XC curves reflects the time-dependent tracer activity and noise level. The XC coefficient for no-MC is lower or similar to the MC XCs. The XC coefficient for no-MC decreases when the tracked motion increases and vice versa for both scans (Fig. E.4 and Fig. E.5) e.g. in the baseline scan there are dips at 38 min and 43 min in the no-MC curve which correspond to the peak motions in Fig. E.4B. This suggests that there is a correlation with the head motion and thus the XC measure and the blurriness of the images.

ROI analysis. Figure E.6 shows the TACs of the left putamen for the three reconstruction schemes applied. The no-MC sequence exhibits irregular patterns during periods of non-exogenous stimuli. Furthermore, the TAC for no-MC has a consistently lower value than the motion-corrected TACs, except for two peaks at 38 min and 43 min where there was registered motion of 4–6 mm.

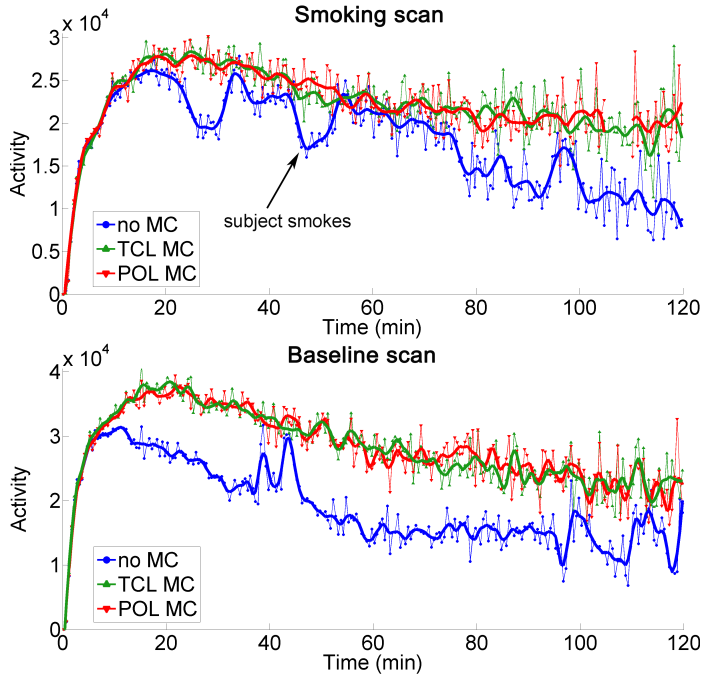


Figure E.6: Time-activity-curves from the left putamen for the 30 s frame reconstruction of the smoking scan (top) and baseline scan (bottom). Blue, green and red curves represent the three reconstructions based on no motion input, the Tracoline system, and the Polaris system, respectively. The solid lines are smoothed curves from the filtering explained in Section E.2.8.

The motion-corrected TACs are significantly different from no-MC (paired t-test, $p < 10^{-38}$ for both TCL and POL tested against no-MC in either smoking or baseline scans).

Overall, there is no significant difference between the TCL MC and POL MC TACs in either scans (paired t-test, $p = 0.28$ and $p = 0.88$ for smoking and baseline scans, respectively). The mean and SD of the percentage TCL and POL difference were: $\mu = 0.47\%$ and SD 10.6% for the smoking scan and $\mu = 0.64\%$ and SD 10.5% for the baseline scan estimated from the dense sample points (SDs with respect to the filtered curve were 4.96% and 5.80% for the smoking and baseline scans, respectively).

Mahalonobis distance. The assumed Poisson type distributions were approximated to normal distributions by computing the square-root of the image values of the 10 min PET images.

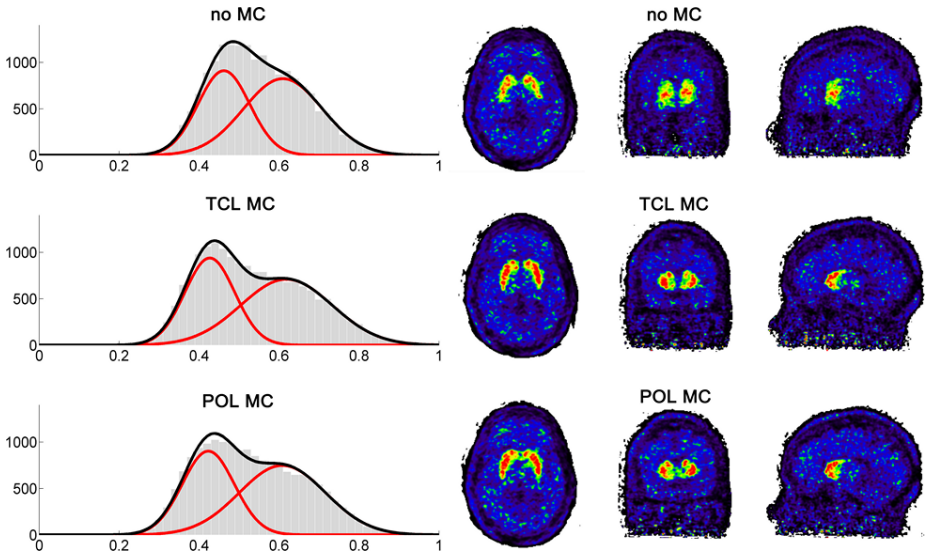


Figure E.7: Left: Histograms from PET image volumes of a 10 min frame at 40–50 min for the smoking scan. A square-root transformation was applied to the data. In this way, the histograms can be approximated by a mixture of Gaussians. The image volumes were masked to a volume containing the high uptake region. The red curves show the results of separation into two Gaussian populations. The black curve is the sum of the two red curves. In the motion-corrected cases better separation of the high uptake and surrounding regions is achieved in comparison to the no motion correction case. Right: PET images of the 10 min frame at 40–50 min for the smoking scan. The contrast is enhanced on the motion-corrected images (row 2–3) compared to the no-MC in row 1.

The images were masked and two normal distributions fitted to the total image distribution. The Gaussian mixture model was fitted using the expectation-maximization algorithm [99]. An example where there is significant motion due to the smoking can be seen in Fig. E.7 (red curves) for a 10 min frame at 40–50 min. The sum of the two separated distributions (black curve) also fits the original image histogram (gray) on the frames not shown. The associated PET images are shown to the right in Fig. E.7. The contrast differences are less expressed on the no-MC images while the caudate nucleus and the putamen regions appear clearer on the motion-corrected PET images (row 2 and 3).

The MD is plotted as a function of time in 10 min frames from 10–60 min (2×5 frames) in Fig. E.8. The first and last frames of the two-hour acquisitions are excluded from the MD measure. This was done due to the high uptake of also non-striatal regions just after tracer injection while the later frames were excluded due to poor image-statistics. The MD for no-MC is lower overall than for the motion-corrected MDs for the smoking scan with major intra-frame motion cf. Fig. E.4A.

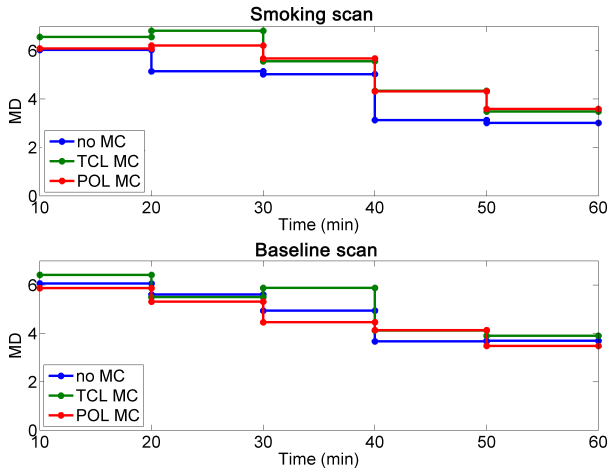


Figure E.8: The Mahalonobis distance (MD) of the PET images with 10 min framing. Top: smoking scan. Bottom: baseline scan. No-MC is significantly lower than TCL MC and POL MC (cf. two-sided ANOVA analysis, $p = 4{\cdot}10^{-13}$) and TCL MC is significantly higher than POL MC (cf. paired t-test, $p = 0.044$) in both scans.

E.3.1 Phantom

The Tracoline and Polaris systems were each used to record motion during a 5 min PET scan of a mannequin head. A test-point 150 mm along the image axes from the center of the image volume was transformed with the MOLAR motion inputs, $(x, y, z) = (150, 150, 150)$. Figure E.9A shows the magnitude of the motion of the test-point, and the bottom plot shows the 3D distances between the two tracking systems. The test-point was moved up to 40 mm including rotation around the center up to 15 degrees (not shown). The mean and SD for the detected 3D motion during the moving period were: $\mu_{TCL} = 19.5$ mm and $\mathrm{SD}_{TCL} = 7.18$ mm for Tracoline and $\mu_{POL} = 19.6$ mm and $\mathrm{SD}_{POL} = 7.11$ mm for Polaris (Fig. E.9A). The SDs of the tracked 3D difference in Fig. E.9B during stationary and moving periods were 0.13 mm and 0.90 mm, respectively.

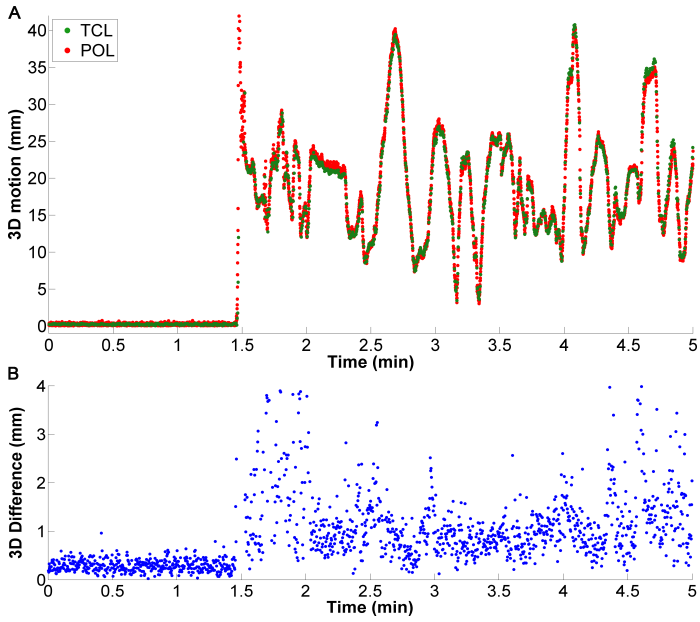


Figure E.9: Comparison of motion registered by the Tracoline (green) and Polaris (red) systems. Top: Estimated motion of a test-point at (x,y,z) = (150, 150, 150) mm in the HRRT coordinate system (Mean and SD of the 3D motion during active period: the $\mu_{TCL} = 19.5$ mm, $\mathrm{SD}_{TCL} = 7.18$ mm, $\mu_{POL} = 19.6$ mm, and $\mathrm{SD}_{POL} = 7.11$ mm). Bottom: The difference of the 3D distance between the tracked point of Tracoline and Polaris. The mean and SD during stationary and moving periods were $\mu = 0.30$ mm and SD = 0.13 for the stationary period and $\mu = 1.16$ mm and SD = 0.90 mm for the moving period.

MC Evaluation. The reconstructed 1 min frames were compared to the first frame of the reconstruction where the mannequin did not move. A visualization of the motion correction for the frame with the largest motion is shown in Fig. E.10. Each row represents a reconstruction method. The reference frame is red and the compared frame is green, thus overlapping regions appear yellow. The last two rows in Fig. E.10 show the results of the MC. Both the TCL MC and POL MC corrected the LORs to the reference position so that the line-sources appear yellow everywhere. All frames were evaluated quantitatively (Fig. E.11) by examining the mutual information (MI) [97, 153] coefficient and the XC with the reference frame.

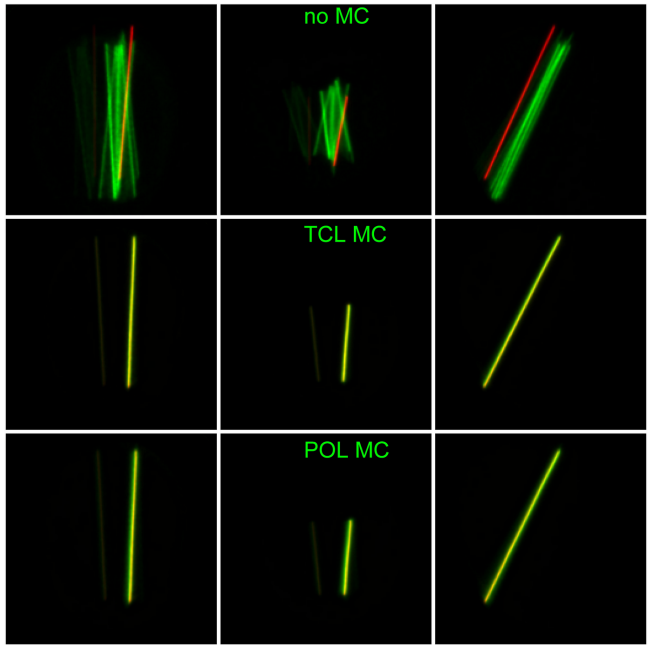


Figure E.10: PET images of frame five from 4–5 min of the phantom study. Display represents image of the line sources embedded in the mannequin head. One of the line sources is hardly seen due to much lower activity then of the other line source (2.2 MBq vs. 17.8 MBq). Each row represents one reconstruction: (1) no-MC, (2) PC, and (3) POL MC. One row of images show the line sources summed along the three image axes: (left) z: axial axis (center) y: anterior/posterior (right) x: left/right. All three axes are 208 mm. The reference frame (frame 1, red) is shown together with the compared frame (frame 5, green). Overlapping pixels are seen as yellow.

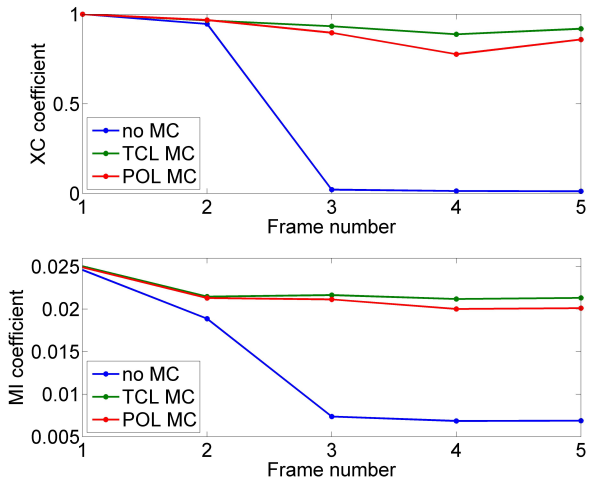


Figure E.11: Similarity results between a reference frame (frame 1, motionless) and a later frame for the phantom study. Tracoline in green and Polaris in red, and no MC in blue. Top plot shows the mutual information (MI) as a function of frame and the bottom plot shows the cross-correlation (XC) measure as a function of frame.

E.4 Discussion

We have presented the first markerless tracking system for motion correction in PET brain imaging. The proposed Tracoline system was tested against a commercial optical tracking system popular for head tracking in PET brain imaging. In previous studies, external tracking systems have not been evaluated against other external tracking devices. This has left the question of whether or not these systems work in human studies, where no ground truth exists, unanswered. We showed highly similar results using the two tracking methods on two human scans and one phantom study.

Concept: The evaluation of the motion correction schemes was based on three measures, purposely leaving out kinetic modeling of the tracer to focus on the effect of motion. Two of the measures, XC and TAC evaluations, showed the time-varying effect of motion on short 30 s time-frames while the third measure, MD, evaluated the effect of MC within the frames (10 min). The XC and TAC measures analyzed a mean motion during the 30 s time-frame. In contrast, the MD measure evaluated the effect of motion correction, under the assumption that the image could be described by 2 activity levels, without discretizing the data into time-frames representing one motion.

Short frames can be problematic for iterative reconstructions but the results shown here support that they can be used for MC evaluation. Cross-correlation was shown previously to be a useful MC evaluation method for ^{18}F-FDG studies [83]. We applied the XC method on short 30 s frames from ^{11}C-racolopride scans. Such short frame durations are normally not recommended since the iterative reconstruction process can converge toward a bias [17], although this effect has not been seen with MOLAR [124]. However, in this study the goal was to evaluate recorded subject motion and its influence on PET reconstruction, so short time-frames were necessary to investigate time-variations despite the potential for introducing small quantitative biases.

Tracoline-based MC: The phantom study demonstrates that the MC scheme based on the Tracoline system functions successfully.

In the phantom experiment, tracking conditions were optimal for both the Tracoline and the Polaris systems since 1) there were no facial movements of the mannequin head AND 2) the Polaris reference tool was firmly fixed. Furthermore, the time-variation in radiotracer distribution was eliminated. The PET images were of high quality, and the degeneration due to motion was quantifiable. We demonstrated that the Tracoline system performed as well or better than the Polaris system (Fig. E.11). An important observation is that the tracking frequency of 5 Hz is adequate for measuring continuous motion, at least for motions as fast as those generally seen in PET brain imaging.

Based on the phantom study, we can conclude that the PET images are improved if correct head motions are given as input to the correction algorithm.

The Tracoline and Polaris systems also agreed overall for the human scans: (1) the SDs of the 3D distance between tracked striatal motions were 0.89 mm and 1.04 mm for the smoking scan and baseline scans, respectively (Fig. E.4); (2) A difference between the TACs of Fig. E.6 for TCL MC and POL MC could not be demonstrated (paired t-test, $p = 0.28$ for smoking scan and $p = 0.88$ for baseline scan). The SD of percentage difference of TCL and POL TAC were limited to 5–11% for the two scans; and (3) the XC coefficients for TCL MC and POL MC were similar. We have demonstrated that the Tracoline MC can be used for human PET studies, the processing and acquisition pipeline works in practice, and the tracking results do not deviate from the state-of-the-art Polaris results.

Reliability of the presented MC schemes: A very powerful innovation of this work is the use of an established external tracking device (Polaris) to evaluate a new external tracking method (Tracoline). Previous studies have used a single external MT method to compare motion correction strategies e.g. [35, 102]. However, in human PET studies evaluation of the tracking performance

requires alternative methods. Evaluation against data-driven methods is not optimal as they cannot be considered as ground truth. Data-driven methods are not independent of the time-varying radiotracer distribution. Hence, motion could be confounded with changes in tracer distribution.

The only major difference between reconstructions with motion data from the Tracoline and Polaris systems is late in the smoking scan (i.e., 105 min after time of injection). At this time data were not registered by the Polaris camera system due to obscuration of marker(s). Overall, data loss such as this is an unusual event. Over the last 100 subjects, ∼0.3% of data is lost with the Polaris.

Advantages of external tracking (and effect of motion on ROI analysis): In this study we present the results of two scans of the same subject with the same protocol except for a smoking period starting at 45 min after time of injection. Thus, the pharmacokinetic model can be assumed to be the same during the first part of each scan (before the smoking period), day-to-day uncertainties notwithstanding. However, visually there is no substantial effect of smoking on the TACs since the motion-corrected TACs of the baseline scan and the smoking scan have similar paths. Thus, the difference we see between TACs of no-MC and motion-corrected TACs is presumed to be mainly due to motion. This is supported by the strong correlation between the registered motion (Fig. E.4A and E.4B) within the left putamen and the uncorrected TACs for the left putamen (Fig. E.6A and E.6B). Note: to detect effects of smoking-induced dopamine release on the ^{11}C-raclopride TAC more advanced kinetic modeling beyond the scope of this study is required.

Proper motion compensation is important for dynamic analysis, as seen in the TACs in Fig. E.6 (where a large difference can be seen between no-MC TACs and motion-corrected TACs). The TAC of no-MC for the baseline scan is dominated by an overall bias compared to the motion-corrected reconstructions. This is mainly expected to be due to the fact that the ROI is placed based on the early images. In addition, a large motion (15–25 mm) between the TX and EM acquisitions which can lead to incorrect scatter and attenuation correction [7]. It is assumed that the patient does not move during the 6 min TX scan. Using an external Tracking system it is, however, possible to record the head motion during the TX scan and discard the scan if too much motion has occurred.

Tracoline versus Polaris: We proposed a method to compare the effect of MC frame-by-frame based on how well striatal ROIs were separated from background using the Mahalonobis distance (MD). The two-population model fit nicely to the image histograms as seen in Fig. E.7. The MD result supports the XC and TAC results. Disagreement of the MD measure between the reconstruction methods was found in frames with large intra-frame motion e.g. the baseline-scan at 30–40 min and 40–50 min.

The MD for TCL MC is just significantly higher than POL MC (paired t-test, $p = 0.044$) suggesting that the Tracoline system performs better than the Polaris system during the 100 min period used for testing. However no broad conclusion can be drawn from only two scans.

This is a proof-of-concept study with a complex design, including two simultaneously running tracking devices. The results are promising and it is presumed that a vision system in-built to a scanner system could provide more features than tracking to the PET acquisition e.g. positioning and visual monitoring of the patients.

Facial movement can potentially reduce the accuracy of the head position estimation with the Tracoline system. This might also be the case for a marker-based tracking system where motion not related to the skull can displace the markers. The camera-based Tracoline system provides information of the facial motion from the raw image-video stream. From the Tracoline video stream the facial motion can be extracted from the overall rigid 3D head motion e.g. from feature extraction or using a template model of the head [104].

E.5 Conclusion

A structured light 3D surface tracking system was demonstrated for markerless head tracking in PET brain imaging. The custom built system was tested against a commercial marker-based system integrated with the HRRT PET scanner both running simultaneously during the PET acquisitions. The tracking recordings of the two systems agreed and the motion correction significantly affected the ROI analysis of the dynamic PET studies and visually improved in contrast recovery of small structures.

Acknowledgements The authors would like to thank Siemens Healthcare A/S and the Danish Agency for Science, Technology and Innovation, which funded the work. We would also like to thank Xiao Jin for technical assistance and the staff of the Yale PET Center.

Appendix F

Technical Report

Evaluation of eye safety issues regarding a near infrared structured light projector system for head pose estimation during PET scanning. By M.Sc.(Eng.) in medicine and technology Oline V. Olesen

F.1 Eye Safety Guidelines

The near infrared (NIR) structured light (SL) system should comply with the lamp safety standard (IEC-62471) [2] or the laser standard (IEC-60825) [3] from the International Electrotechnical Commission (IEC). The Application Note entitled: "Eye Safety of IREDs used in Lamp Applications" from OSRAM [5] is based on the lamp standard IEC-62471, and this covers most light emitting diode (LED) applications. The "American National Standards for Safe Use of Lasers" from the Laser Institute of America (ANSI) provide guidance for the safe use of laser systems [1]. The following limits for exposure are based on the mentioned OSRAM note and ANSI guidance, assuming the IEC standards are similar. The risk assessment and classification is to be done with the final product and not the built-in components. In this case the combined near infrared emitting diode (NIRED) and the pico projector. Figure F.1 show the NIR SL system, which consists of two CCD cameras and a modified miniaturized light projector. One of the original LEDs of the projector has been replaced with a NIRED working as the irradiating part of the NIR SL system, Tracoline.

When irradiating the human body there are three hazard exposure limits to consider: (1) for the cornea, (2) thermal hazard exposure limit for skin ($t < 10$ s), and (3) thermal hazard exposure limit for the retina. In our case the time

Figure F.1: The NIR SL system consists of two CCD cameras and a modified nonvisible light projector located in the center. Right: The application of the NIR SL system where it is seen attached to the Siemens HRRT PET scanner for head pose registration during the PET acquisition.

for exposure is above 10 s and in this case the thermal exposure limit for skin is always obeyed if the thermal hazard exposure limit for the retina is obeyed.

The maximum permissible exposure (MPE) or irradiance limit for the cornea is fixed when the exposure duration exceeds 1000 s [5] as shown in Eq. (F.1) which refers to the IEC-62471 standard [5].

$$MPE_{\text{cornea}} = \sum_{\lambda=850}^{3000} E_\lambda \Delta\lambda \leq 100 \frac{\text{W}}{\text{m}^2} \quad \text{for} \quad t > 1000 \text{ s} \tag{F.1}$$

where E_λ [W/(m^{2}nm)] is the spectral irradiance, $\Delta\lambda$ [nm] is the spectral bandwidth.

The exposure limit for the retina is estimated from the ANSI guidance for extended sources (non-point sources) with wavelengths from 750 nm to 1050 nm and exposure duration up to 500 minutes (8.3 hours) [1, p. 62, 75, and Tb. 5b]. The MPE for the retina was calculated as:

$$MPE_{\text{retina}} = 1.8 C_A C_E T_2^{-0.25} 10^{-3} \quad [\frac{\text{W}}{cm^2}] \tag{F.2}$$

where the correction factors C_A, C_E, and T_2 are given as listed in Table F.1 [1, p. 76 Tb. 6].

The angular subtense α (see Fig. F.2) varies as a function of the distance to the projector d, as does the time parameter T_2, as shown in Fig. F.3(a) and Fig. F.3(b) respectively. Figure F.3(c) shows the MPE for the retina as a function of d based on Eq. (F.2) with the parameters given in Table F.1. E.g. the

Table F.1: Parameters used in Eq. (F.2). All wavelengths must be in μm and all angles in mrad for the calculations.

Parameters		Estimation	Criteria
α [rad]	Apparent angle subtended by a source at the location of the viewer	$^{*}\alpha = \arctan(r/d)$	λ =[750 1050] nm
C_A	Wavelength correction factor	$^{**}C_A = 10^{2(\lambda - 0.700)}$	λ =[750 1050] nm
C_E	Extended source correction factor	$^{***}C_E = \alpha/\alpha_{min}$	$\alpha_{min} < \alpha < \alpha_{max}$
T_2 [s]	Exposure duration beyond which the thermal MPE is constant in terms of irradiance	$10 \cdot 10^{(\alpha - 1.5)/98.5}$ 100 s	$\alpha_{min} < \alpha < \alpha_{max}$ $\alpha > 100$ mrad

$^{*}d$:distance from the viewer to the LED. r: source size of the active field-of-view of the projector (See Fig. F.2).

$^{**}\lambda$: central wavelength of the NIR LED [μm].

$^{***}\alpha_{min} = 1.5$ mrad and $\alpha_{max} = 100$ mrad for $\lambda = [400\ 1400]$ nm.

MPE for the retina at 100 mm is calculated as:

$$MPE_{\text{retina}} = 1.8 \cdot 10^{2(0.850\mu\text{m} - 0.700)} \frac{\arctan(\frac{7\text{mm}}{100\text{mm}})10^3}{15\text{mrad}} \ldots \cdot \left(10 \cdot 10^{\frac{\arctan(\frac{7\text{mm}}{100\text{mm}})10^3 - 1.5}{98.5}}\right)^{-0.25} 10^{-3} \cdot 10^4 \frac{cm^2}{\text{m}^2} = \underline{631 \frac{\text{W}}{\text{m}^2}}$$

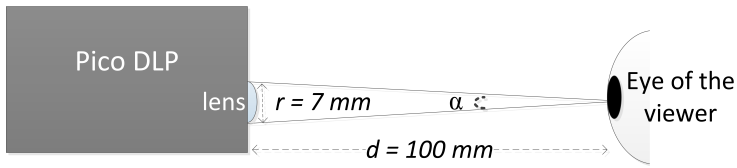


Figure F.2: Drawing of the apparent visual angle α as calculated from the source size r of the illuminating region of the projector and the distance from the viewer d.

F.2 Characterization of the NIR SL System

The NIR SL system's illumination is uniformly distributed over the area covered by the projector, thus the intensity is given by power/area as a function of the distance to the projector. The area covered by the projector as a function of distance is presented in Fig. F.4(a) The exposure is highest closest to the DLP and for the NIR SL system this is at a distance of 100 mm from the lens. The exposed area at 100 mm is 38 mm $\times$ 56 mm = 2128 mm^2.

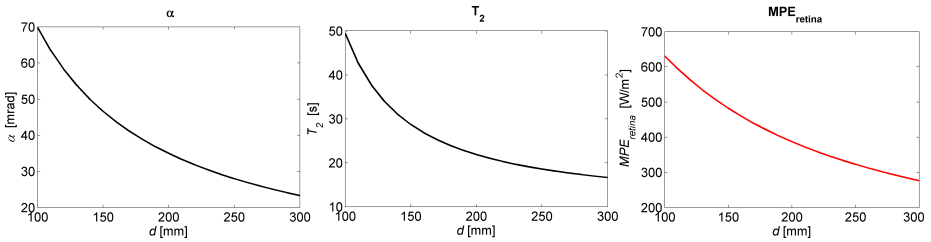


Figure F.3: (a) The angular subtense. (b) The exposure duration beyond which the thermal MPE is constant in terms of irradiance. (c) The maximum permissible exposure (MPE) for the retina. All three parameters are shown as a function of the distance to the projector d.

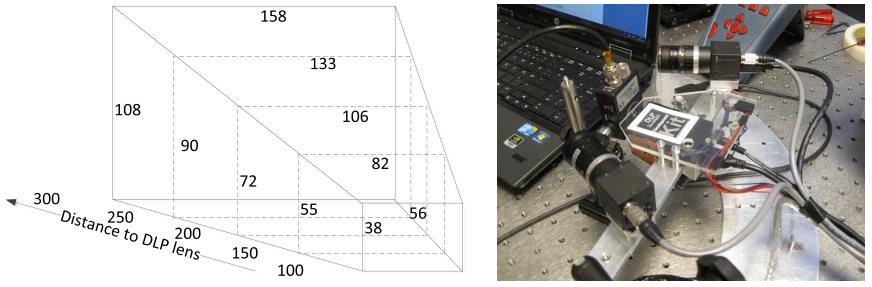


Figure F.4: (a) Measuring volume of the NIR SL system [mm]. (b) Measurement of the NIRED beam power through the optics of the projector with a photodiode placed in front of the projector lens.

The beam power of the NIR SL system has been measured by PhD student MSc (Physics) Otto Nielsen. The power output was measured as a function of the NIRED's input current from 0–1 A with a power meter using a CCD detector. The detector was placed as close as possible to the projector lens, less than 1 cm from the lens. The current of the NIRED was manually set on the power supply. Figure F.5 shows the power dependence that was measured twice during the characterization of the combined diode and projector. The projector was projecting an 8-bit white image (B/W value = 255) with 60 Hz which produced maximum output intensity and thus worked as a worse-case estimate. The maximum measured power of the combined NIRED and projector was 6.2 mW at 1 A.

The light source in the NIR SL system is an 850 nm NIRED with a spectral band of 775–900 nm. The maximum specified radiation flux of the NIRED is 440 mW for exposure pulses of 100 μs and a maximum input current of 1 A. The actual power of NIR SL systems was around 70 times less than the specification of the NIRED, which was expected since the optics of the projector are not optimized for NIR use. Further, the NIR SL system is non-pulsed and pulsed irradiation increases the maximum power.

The power-time profile was measured with a photodiode using an InGaAs detector (DET10C, Thorlabs) (see Fig. F.4(b)). Figure F.6 shows the time structure of the projector output beam for three images with different 8-bit gray-scaled values of 32, 128, and 255 respectively. The beam is periodic with 16.7 ms (60 Hz) for the two nonwhite images (blue and green profiles). The profiles show how the DMDs (micro mirror devices) are switched on and off during a cycle of 16.7 ms. When the DMDs are on, the beam power is positive and goes up to 5.3 V. It should be noted that the amplitude is a relative measure in volts that could be converted to watts using a calibration coefficient. The time the DMDs are switched on increases with the gray value of the projected image, and for the white image the power is constant.

It is important to notice that there is no high peaks exceeding the mean beam power when projecting a white image. Thus the measured mean power as a function of the current of a white image is the maximum power of the NIR SL system, as shown in Fig. F.5.

The actual exposure E_{actual} is power P per area A and the maximum exposure for the NIR SL system within the measuring volume shown in Fig. F.4(a) is at d =100 mm with a max exposure of:

$$E_{\text{actual}} = \frac{P}{A} = \frac{6.2 \cdot 10^{-8}\text{W}}{2.13 \cdot 10^{-8}\text{m}^2} = \underline{2.9\frac{\text{W}}{\text{m}^2}} \tag{F.3}$$

Figure F.7(a) shows the actual exposure as a function of the distance to the projector and Fig. F.7(b) shows the two MPE for the cornea and the retina and the actual exposure of the NIR SL system. The actual exposure is below the MPEs for all distances within the measuring volume of the NIR SL system. A quantitative visualization of how much the actual exposure is lower than the MPEs is shown in Fig. F.7(c) where the ratio between the actual exposure and the MPEs are seen as a function of d. The ratio is smallest closest to the projector and the actual exposure is a factor of 34 less than the MPE for the cornea and a factor of 217 less than the MPE for the retina at $d = 100$ mm (Fig. F.7(c) blue and red curves at $d = 100$ mm).

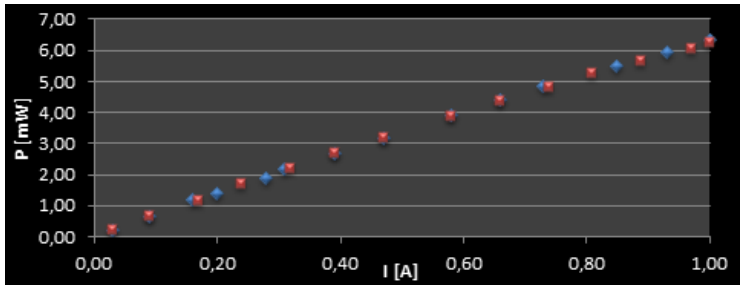


Figure F.5: Measured beam power as a function of the current when projecting a white image (B/W level = 255) repeated twice indicated by blue and red markers.

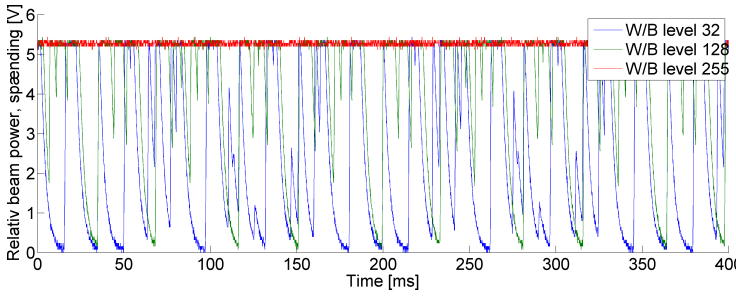


Figure F.6: Beam power as a function of time with a constant current of 0.7 A. Three different 8-bit gray scaled images were projected: (Blue) dark gray image with B/W level of 32. (Green) gray image with intensity of 128. (Red) white image with B/W level of 255.

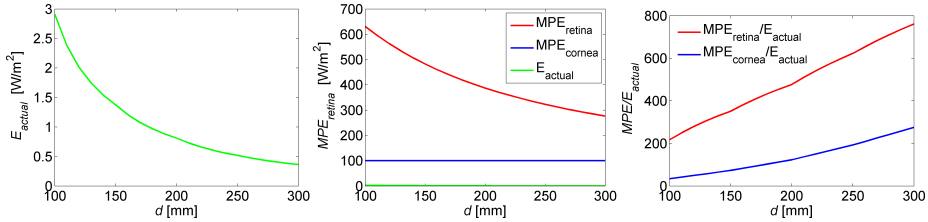


Figure F.7: (a) The actual exposure E_{actual} of the NIR SL system. (b) The maximum permissible exposure (MPE) for the retina and the cornea and E_{actual}. (c) The ratio between the MPEs and E_{actual}. All three plots are shown as a function of the distance to the projector d.

F.3 Conclusion

The maximum exposure of the NIR SL system is 2.9 W/m^2 at the minimum measuring distance at 100 mm. That is 34 and 217 times below the MPE of the cornea (100 W/m^2) and the retina (631 W/m^2) respectively. The NIR SL system does not exceed the MPE for the eyes in the worst-case scenario when projecting a constant white image 10 cm away for up to 8 hours according to the OSRAM note based on the IEC-62471 standard and the American National Standards for safe use of lasers. The worst-case scenario is above the actual application of the NIR SL system where the projected images are phase-shifted cosine patterns with intensities lower than a white image and also varying between dark and light intensities. Furthermore, the distance to the eyes is more likely to be 15 cm and the eyes will normally be closed during the scanning. I consider the NIR SL system to be safe to use for research for the tracking application with the HRRT PET scanner described earlier. The evaluation is based on the American National Standards and it is assumed that international standards are in line with these. If the system is to be used in the clinical routine and be commercially available, the system should be classified by an external company.

Appendix G

System Classification

Classification of the optics of the Tracoline system from an external company: Dansk Fundamental Metrologi A/S, Denmark by PhD Jan C. Petersen

Dansk Fundamental Metrologi A/S
Matematiktorvet 307
DK-2800 Kgs. Lyngby
Tel +45 4593 1144
www.dfm.dtu.dk

Certifikat nr. PM1101
Side 1 af 2

Calibration certifikat

Optical power

Client	Rigshospitalet
Address	Blegdamsvej 9, KF4011 2100 København Ø
Telephone/Fax	35 45 16 77
Contact person	Professor Liselotte Højgaard
Date received	2011-03-28
Date performed	2011-03-28

Item	Near-Infrared light projector system with an OSRAM 850 nm LED (nominal 440 mW) and the original lens (optimized for the visible wavelength region)
Identification	Home built system
Serial no.	NA

The calibration results are traceable to the Danish primary standard.

If the calibration certificate is quoted partially, DFM must give written consent.

Dato: 2011-04-01

Jan C. Petersen
Ph.D.

DANAK akkreditering 255
Dansk Fundamental Metrologi A/S
Matematiktorvet 307, DK-2800 Kgs. Lyngby

Method

The power calibration is performed according to DFM procedure Q2KAL402.

The working standard used is the thermopile T540. The working standard was most recently calibrated 2010-10-23.

The total power was measured approximately 10 mm from the source.

The calibrations were performed at a room temperature of 22 °C ± 1 °C.

The reported expanded uncertainty of measurement is stated as the standard uncertainty multiplied by a coverage factor $k = 2$. For a normal distribution this corresponds to a coverage probability of approximately 95%. The standard uncertainty of measurement has been determined in accordance with EA publication EAL-R2.

Calibration parameters		*Results*	
Wavelength	Current	Power	Uncertainty
nm	A	mW	mW
855	1	6,43	±0,14

Classification

The classification is partly based on the following statements in a report by Oline V. Olesen:

It is important to notice that there is no high peaks exceeding the mean beam power when projection a white image. Thus the measured power as a function of the current of a white image is the maximum power of the NIR SL system as shown in Figure 5.

The exposed area at 100 nm is 38 mm × 56 mm = 2128 mm².

The uncertainty associated with the exposed area is estimated to be ±241 mm²

The standard EN 60825-1: 2007 (Safety of laser products – Part 1: Equipment classification and requirements) has been used to calculate the maximum permissible exposure (MPE) and the accessible emission limit (AEL). The latter indicates the maximum accessible emission permitted within a particular class.

MPE (Table A.2) = (639 ± 81) W/m² at a distance of 100 mm from the source.

The irradiance of the source at a distance of 100 mm from the source is: (3,02 ± 0,44) W/m²

AEL (Table 5) = (25 ± 3) mW. This value is the AEL for a Class 1 laser product with a configuration as described in the above-mentioned report and viewed at a distance larger than 100 mm with the naked (unaided) eye.

A class 1 laser product is any laser product, which during operation does not permit human access to accessible laser radiation in excess of the accessible emission limits of Class 1 for applicable wavelengths and emission durations.

Bibliography

[1] *American National Standards for Safe Use af Lasers.* American National Standards Institute (ANSI), Laser Insitute of America.

[2] *The lamp safety standard (IEC-62471).* The International Electrotechnical Commission (IEC).

[3] *The laser standard (IEC-60825).* The International Electrotechnical Commission (IEC).

[4] Polaris vicra user guide. *Northern Digit Inc., Canada*, Revision 2, 2005.

[5] *Eye Safety of IREDs used in Lamp Applications.* OSRAM Opto Semiconductors, 2010.

[6] S. Anishchenko, V. Osinov, D. Shaposhnikov, L. Podlachikova, R. Comley, and X.W. Gao. Toward a robust system to monitor head motions during PET based on facial landmark detection: A new approach. *IEEE International Symposium on Computer-Based Medical Systems Conference Record*, 50–52, 2008.

[7] J.M. Anton-Rodriguez, M. Sibomana, M.D. Walker, M.C. Huisman, J.C. Matthews, M. Feldmann, S.H. Keller, and M.C. Asselin. Investigation of motion induced errors in scatter correction for the HRRT brain scanner. *IEEE Nuclear Science Symposium Conference Record*, pages 2935–2940, 2010.

[8] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate (corresp.). *IEEE Transactions on Information Theory*, 20(2):284–287, 1974.

[9] P. Bahl and V.N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. *Proceedings of IEEE Computer and Communications Societies*, 2:775–784, 2000.

[10] B. Bai, Q. Li, CH Holdsworth, E. Asma, YC Tai, A. Chatziioannou, and RM Leahy. Model-based normalization for iterative 3D PET image reconstruction. *Physics in medicine and biology*, 47:2773, 2002.

[11] D. L. Bailey, D. W. Townsend, P. E. Val, and M. N. Maisey. *Position Emission Tomography*. Springer, 2005.

[12] R.D. Beach, H. Depold, G. Boening, P.P. Bruyant, B. Feng, H.C. Gifford, M.A. Gennert, S. Nadella, and M.A. King. An adaptive approach to decomposing patient-motion tracking data acquired during cardiac SPECT imaging. *IEEE Transactions on Nuclear Science*, 54(1):130–139, 2007.

[13] P.J. Besl and N.D. McKay. A method of registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[14] P. Bühler, U. Just, E. Will, J. Kotzerke, and J.V.D Hoff. An accurate method for correction of head movement in PET. *IEEE Transactions on Medical Imaging*, 23:1176–1185, 2004.

[15] F. Bitte, G. Dussler, and T. Pfeifer. 3D micro-inspection goes DMD. *Optics and lasers in engineering*, 36(2):155–167, 2001.

[16] P.M. Bloomfield, T.J. Spinks, J. Reed, L. Schnorr, A.M. Westrip, L. Livieratos, R. Fulton, and T. Jones. The design and implementation of a motion correction scheme for neurological PET. *Physics in Medicine and Biology*, 48(8):959–978, 2003.

[17] R. Boellaard, A. van Lingen, and A.A. Lammertsma. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *Journal of Nuclear Medicine*, 42(5):808, 2001.

[18] J.Y. Bouguet. Camera calibration toolbox for matlab. *http://www.vision.caltech.edu/bouguetj/calib_doc/*, June, 2008.

[19] J.H. Bruning, D.R. Herriott, J.E. Gallagher, D.P. Rosenfeld, A.D. White, and D.J. Brangaccio. Digital wavefront measuring interferometer for testing optical surfaces and lenses. *Applied Optics*, 13(11):2693–2703, 1974.

[20] Q. Cai and J.K. Aggarwal. Tracking human motion using multiple cameras. *Proceedings of the 13th International Pattern Recognition*, 3:68–72, 1996.

[21] R.E. Carson. Tracer kinetic parametric imaging in PET. *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 611–615, 2004.

[22] R.E. Carson, W.C. Barker, J.-S. Liow, and C.A. Johnson. Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT. *IEEE Nuclear Science Symposium Conference Record*, 5:3281–3285, 2003.

[23] C. Catana, T. Benner, A. van der Kouwe, L. Byars, M. Hamm, D.B. Chonde, C.J. Michel, G. El Fakhri, M. Schmand, and A.G. Sorensen. MRI-assisted PET motion correction for neurologic studies in an integrated MR-PET scanner. *Journal of Nuclear Medicine*, 52(1):154, 2011.

[24] F. Chen, G.M. Brown, and M. Song. Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39:10, 2000.

[25] W.S. Cleveland and S.J. Devlin. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 596–610, 1988.

[26] C. Comtat, F. C. Sureau, and M. Sibomana. Image based resolution modeling for the HRRT OSEM reconstructions software. *IEEE Nuclear Science Symposium Conference Record*, 4120–4123, 2008.

[27] N. Costes, A. Dagher, K. Larcher, A.C. Evans, D.L. Collins, and A. Reilhac. Motion correction of multi-frame PET data in neuroreceptor mapping: Simulation based validation. *Neuroimage*, 47(4):1496–1505, 2009.

[28] K. Creath. Phase-measurement interferometry techniques. *Progress in optics*, 26:349–393, 1988.

[29] M.E. Daube-Witherspoon, Y.C. Yan, M.V. Green, R.E. Carson, K.M. Kempner, and P. Herscovitch. Correction for motion distortion in PET by dynamic monitoring of patient position. *Journal of Nuclear Medicine*, 31(5):816, 1990.

[30] H.W.A.M. De Jong, F.H.P. van Velden, R.W. Kloet, F.L. Buijs, R. Boellaard, and A.A. Lammertsma. Performance evaluation of the ECAT HRRT: an LSO-LYSO double layer high resolution, high sensitivity scanner. *Physics in Medicine and Biology*, 52(5):1505–26, 2007.

[31] G. Delso and S. Ziegler. PET/MRI system design. *European journal of nuclear medicine and molecular imaging*, 36:86–92, 2009.

[32] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, 39:1–38, 1977.

[33] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[34] K. Dinelle, S. Blinder, C. Ju-Chieh, S. Lidstone, K. Buckley, T.J. Ruth, and V. Sossi. Investigation of subject motion encountered during a typical positron emission tomography scan. *IEEE Nuclear Science Symposium Conference Record*, 6:3283–3287, 2006.

[35] K. Dinelle, H. Ngo, S. Blinder, N. Vafai, G. Topping, and V. Sossi. Frame-to-frame image realignment assessment tool for dynamic brain positron emission tomography. *Medical Physics*, 38:773, 2011.

[36] J. Ehrhardt, R. Werner, T. Frenzel, D. Säring, W. Lu, D. Low, and H. Handels. Reconstruction of 4D-CT data sets acquired during free breathing for the analysis of respiratory motion. *Proceedings of SPIE - The International Society for Optical Engineering, Medical Imaging*, 6144:614414–1–8, 2006.

[37] P. Enge and P. Misra. Special issue on global positioning system. *Proceedings of the IEEE*, 87(1):3–15, 1999.

[38] D. Essig, S. Schmid, M. Dawood, X. Jiang, and K. P. Schaefers. Camera-based tracking system for freely moving mice inside a small animal PET scanner. *IEEE Nuclear Science Symposium Conference Record*, pages MIC18.M–34, 2011.

[39] T.L. Faber, N. Raghunath, D. Tudorascu, and J.R. Votaw. Motion correction of PET brain images through deconvolution: I. theoretical development and analysis in software simulations. *Physics in Medicine and Biology*, 54(3):797–811, 2009.

[40] B. Fischl, D.H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, et al. Whole brain segmentation:: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.

[41] J. Forma, U. Tuna, and U. Ruotsalainen. Estimation of rigid body motion parameters for the ECAT HRRT data without image reconstruction. *IEEE Nuclear Science Symposium Conference Record*, 3534–3540, 2010.

[42] C. Forman, M. Aksoy, J. Hornegger, and R. Bammer. Self-encoded marker for optical prospective head motion correction in MRI. *Medical Image Analysis*, 15(5):708–719, 2011.

[43] G. Frankowski and R. Hainich. DLP-based 3D metrology by structured light or projected fringe technology for life sciences and industrial metrology. *Proceedings of SPIE - The International Society for Optical Engineering*, 7210:72100C, 2009.

[44] R. Fulton, L. Tellmann, U. Pietrzyk, O. Winz, I. Stangier, I. Nickel, A. Schmid, S. Meikle, and H. Herzog. Accuracy of motion correction methods for PET brain imaging. *IEEE Nuclear Science Symposium Conference Record*, 7:4226–4230, 2004.

[45] R.R. Fulton, S.R. Meikle, S. Eberl, J. Pfeiffer, C.J. Constable, and M.J Fulham. Correction for head movements in positron emission tomography using an optical motion-tracking system. *IEEE Transactions on Nuclear Science*, 49:116–123, 2002.

[46] V. Ganapathi, C. Plagemann, D. Koller, and S. Thrun. Real time motion capture using a single time-of-flight camera. *IEEE Conference on Computer Vision and Pattern Recognition*, 755–762, 2010.

[47] D.M. Gavrila. The visual analysis of human movement: A survey* 1. *Computer vision and image understanding*, 73(1):82–98, 1999.

[48] J. Geng. Structured-light 3D surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2):128–160, 2011.

[49] Z.J. Geng. Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems. *Optical Engineering*, 35:376, 1996.

[50] Z.J. Geng. High speed three dimensional imaging method. US Patent 6,028,672, February 22 2000.

[51] Z.J. Geng. Method and system for three-dimensional imaging using light pattern having multiple sub-patterns. US Patent 6,700,669, March 2 2004.

[52] D.C. Ghiglia, G.A. Mastin, and L.A. Romero. Cellular-automata method for phase unwrapping. *Journal of the Optical Society of America A*, 4(1):267–280, 1987.

[53] J.S. Goddard, S.S. Gleason, M.J. Paulus, R. Kerekes, S. Majewski, V. Popov, M. Smith, A. Weisenberger, B. Welch, and R. Wojcik. Pose measurement and tracking system for motion-correction of unrestrained small animal PET/SPECT imaging. *IEEE Nuclear Science Symposium Conference Record*, 3:1824–1827, 2003.

[54] R. Goldstein, H. Zebker, and C. Werner. Satellite radar interferometry-two-dimensional phase unwrapping. *Radio Science*, 23(4):713–720, 1988.

[55] S.R. Goldstein, M.E. Daube-Witherspoon, M.V. Green, and A. Eidsath. A head motion measurement system suitable for emission computed tomography. *IEEE Transactions on Medical Imaging*, 16(1):17–27, 1997.

[56] R.C. Gonzalez and R.E. Woods. Digital image processing. *Prentice Hall*, pages 205–208, 701–704, 2002.

[57] C. Gramkow. On averaging rotations. *Journal of Mathematical Imaging and Vision*, 15(1):7–16, 2001.

[58] M.V. Green, J. Seidel, S.D. Stein, T.E. Tedder, K.M. Kempner, C. Kertzman, and T.A. Zeffiro. Head movement in normal subjects during simulated PET brain imaging with and without head restraint. *Journal of Nuclear Medicine*, 35(9):1538–1546, 1994.

[59] A.P. Happonen and U. Ruotsalainen. Three-dimensional alignment of scans in a dynamic PET study using sinusoidal trajectory signals of a sinogram. *IEEE Transactions on Nuclear Science*, 51(5):2620–2627, 2004.

[60] R. Hartley and A. Zisserman. Multiple view geometry in computer vision. *Cambridge Univ Pr*, (9), 2000.

[61] R.I. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593, 1997.

[62] K. Herholz, S. Westwood, C. Haense, and G. Dunn. Evaluation of a calibrated 18F-FDG PET score as a biomarker for progression in alzheimer disease and mild cognitive impairment. *Journal of Nuclear Medicine*, 52(8):1218, 2011.

[63] M.A. Herráez, D.R. Burton, M.J. Lalor, and M.A. Gdeisat. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. *Applied Optics*, 41:7437–7444, 2002.

[64] H. Herzog, L. Tellman, R. Fulton, and U. Pietrzyk. Motion correction in PET brain studies. *Proceedings of The Fourth International Workshop on Multidimensional Systems*, 178–181, 2005.

[65] H. Herzog, L. Tellmann, R. Fulton, I. Stangier, E.R. Kops, K. Bente, C. Boy, R. Hurlemann, and U. Pietrzyk. Motion artifact reduction on parametric PET images of neuroreceptor binding. *Journal of Nuclear Medicine*, 46(6):1059, 2005.

[66] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.

[67] J. Hirvonen, J. Johansson, M. Teräs, V. Oikonen, V. Lumme, P. Virsu, A. Roivainen, K. Någren, C. Halldin, L. Farde, and J Hietala. Measurement of striatal and extrastriatal dopamine transporter binding with high-resolution PET and [11-C] PE2I: quantitative modeling and test-retest reproducibility. *Journal of Cerebral Blood Flow and Metabolism*, 28(5):1059–1069, 2008.

[68] S. Holm, P. Toft, and M. Jensen. Estimation of the noise contributions from blank, transmission and emission scans in PET. *IEEE Transactions on Nuclear Science*, 43(4 1):2285–2291, 1996.

[69] I. Hong and A.J. Reader. Ultra fast 4D PET image reconstruction with user-definable temporal basis functions. *IEEE Nuclear Science Symposium Conference Record*, 5475–5478, 2008.

[70] B.K.P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.

[71] M.A. Horton and A.R. Newton. Method and apparatus for determining position and orientation of a moveable object using accelerometers. *US Patent 5,615,132*, March 25 1997.

[72] D. Hu, C. Hayden, M. Casey, and Z. Burbar. Stereo computer vision system for measuring movement of patient's head in PET scanning. *IEEE Nuclear Science Symposium Conference Record*, 5:2864–2867, 2004.

[73] P.S. Huang, Q. Hu, F. Jin, and F.P. Chiang. Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring. *Optical Engineering*, 38(6):1065–1071, 1999.

[74] B.F. Hutton, A.Z. Kyme, Y.H. Lau, D.W. Skerrett, and R.R. Fulton. A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography. *IEEE Transactions on Nuclear Science*, 49(1):188–194, 2002.

[75] Point Grey Research Inc. Imaging products timestamping and different timestamp mechanisms. *Knowledge Base*, Article ID: 99, 2010.

[76] K. Itoh. Analysis of the phase unwrapping algorithm. *Applied Optics*, 21(14):2470, 1982.

[77] M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2):143–156, 2001.

[78] X. Jin, T. Mulnix, B. Planeta-Wilson, J.D. Gallezot, and R.E. Carson. Accuracy of head motion compensation for the HRRT: Comparison of methods. *IEEE Nuclear Science Symposium Conference Record*, 3199–3202, 2009.

[79] S. Johannessen. Time synchronization in a local area network. *IEEE Control Systems Magazine*, 24(2):61–69, 2004.

[80] N. Jojic, J. Gu, H. Shen, and T. Huang. 3-D reconstruction of multipart self-occluding objects. *Computer Vision*, 455–462, 1997.

[81] M.E. Kamasak, C.A. Bouman, E.D. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Transactions on Medical Imaging*, 24(5):636–650, 2005.

[82] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson Surface Reconstruction. *Proceedings of Symposium on Geometry*, 61–70, 2006.

[83] S.H. Keller, M. Sibomana, O.V. Olesen, C. Svarer, S. Holm, F. Andersen, and L. Højgaard. Methods for motion correction evaluation using FDG human brain scans on a high resolution PET scanner. *Journal of Nuclear Medicine*, Accepted for publication.

[84] M. Kjer, O.V. Olesen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Geometric calibration between PET scanner and structured light scanner. *MICCAI Workshop on Mesh Processing in Medical Image Analysis*, 2011.

[85] C. Knöß. *Evaluation and Optimazation of the High Resolution Research Tomograph (HRRT).* PhD thesis, ISBN 3-8322-3337-7, 2004.

[86] C. Knoess, W.-D. Heiss, R. Nutt, J. Rist, C. Michel, Z. Burbar, L. Eriksson, V. Panin, L. Byars, M. Lenox, and K. Wienhard. Evaluation of single photon transmission for the HRRT. *IEEE Nuclear Science Symposium Conference Record*, 3:1936–1940, 2003.

[87] A. Kyme, S. Se, S. Meikle, C. Baldock, W. Ryder, and R. Fulton. Novel SLAM-based markerless motion tracking of conscious unrestrained rodents in PET. *IEEE Nuclear Science Symposium Conference Record*, MIC16-4, 2011.

[88] AZ Kyme, VW Zhou, SR Meikle, and RR Fulton. Real-time 3D motion tracking for small animal brain PET. *Physics in Medicine and Biology*, 53:2651, 2008.

[89] F. Lamare, MJ Ledesma Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C.C. Le Rest, AJ Reader, and D. Visvikis. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in medicine and biology*, 52:5187, 2007.

[90] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal of computer assisted tomography*, 8(2):306–316, 1984.

[91] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE Journal of Quantum Electronics*, 37(3):390–397, 2001.

[92] T. Lindeberg. Scale-space theory in computer vision. *Springer*, 1993.

[93] B.J. Lopresti, A. Russo, W.F. Jones, T. Fisher, D.G. Crouch, D.E. Altenburger, and D.W. Townsend. Implementation and performance of an optical motion tracking system for high resolution brain PET imaging. *IEEE Transactions on Nuclear Science*, 46(6):2059–2067, 1999.

[94] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[95] H.J. Luinge and P.H. Veltink. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical and Biological Engineering and Computing*, 43(2):273–282, 2005.

[96] W.P.T. Ma, G. Hamarneh, G. Mori, K. Dinelle, and V. Sossi. Motion estimation for functional medical imaging studies using a stereo video head pose tracking system. *IEEE Nuclear Science Symposium Conference Record*, 4086–4090, 2008.

[97] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[98] O. Mawlawi, D. Martinez, M. Slifstein, A. Broft, R. Chatterjee, D.R. Hwang, Y. Huang, N. Simpson, K. Ngo, and R. Van Heertum. Imaging human mesolimbic dopamine transmission with positron emission tomography: I. accuracy and precision of D2 receptor parameter measurements in ventral striatum. *Journal of Cerebral Blood Flow & Metabolism*, 21(9):1034–1057, 2001.

[99] G.J. McLachlan and D. Peel. *Finite mixture models*, volume 299. Wiley-Interscience, 2000.

[100] M. Menke, M.S. Atkins, and K.R. Buckley. Compensation methods for head motion detected during PET imaging. *IEEE Transactions on Nuclear Science*, 43(1):310–317, 1996.

[101] T.B. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding*, 104(2-3):90–126, 2006.

[102] A.J. Montgomery, K. Thielemans, M.A. Mehta, F. Turkheimer, S. Mustafovic, and P.M. Grasby. Correction of head movement on PET studies: comparison of methods. *Journal of Nuclear Medicine*, 47(12):1936–1944, 2006.

[103] H. Muraishi, T. Hasegawa, Y. Fukushima, K. Yoda, K. Maruyama, Y. Shiba, M. Muraishi, T. Yamaya, E. Yoshida, and H. Murayama. New tracking method for head motion using a single camera and a solid marker. *Medical Imaging Technology*, 24(4):320–8, 2006.

[104] E. Murphy-Chutorian and M.M. Trivedi. Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):607–626, 2009.

[105] C. Nimsky, O. Ganslandt, S. Cerny, P. Hastreiter, G. Greiner, and R. Fahlbusch. Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging. *Neurosurgery*, 47(5):1070, 2000.

[106] P.J. Noonan, W.A. Hallett, T. Cootes, and R. Hinz. The design and initial calibration of an optical 3D acquisition system for head motion tracking. *IEEE Nuclear Science Symposium Conference Record*, MIC18.M-10, 2011.

[107] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (swissranger TM). *Optical Design and Engineering*, 5249:534–545, 2004.

[108] O.V. Olesen, M.R. Jørgensen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Structured light 3D tracking system for measuring motions in PET brain imaging. *Proceedings of SPIE - The International Society for Optical Engineering, Medical Imaging*, 7625:76250X–11, 2010.

[109] O.V. Olesen, S.H. Keller, M. Sibomana, R Larsen, B. Roed, and L. Højgaard. Automatic thresholding for frame-repositioning using external tracking in PET brain imaging. *IEEE Nuclear Science Symposium Conference Record*, 2669–2675, 2010.

[110] O.V. Olesen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. External motion tracking for brain imaging: Structured light tracking with invisible light. *IEEE Nuclear Science Symposium Conference Record*, 2735–2737, 2010.

[111] O.V. Olesen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Motion tracking in narrow spaces: A structured light approach. *Lecture Notes in Computer Science, Springer*, 6363:253–260, 2010.

[112] O.V. Olesen, R.R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Motion tracking for medical imaging: A non-visible structured light tracking approach. *IEEE Transactions on Medical Imaging*, 2011, to appear, online version available.

[113] O.V. Olesen, R.R. Paulsen, R.R. Jensen, S.H. Keller, M. Sibomana, L. Højgaard, B. Roed, and R. Larsen. 3D surface realignment tracking for medical imaging: A phantom study with PET motion correction. *Image-based Geometric Modeling and Mesh Generation by Springer, E. Yongjie Zhang*, (to appear).

[114] O.V. Olesen, R.R. Paulsen, S.H. Keller, , L. Højgaard, B. Roed, and R. Larsen. Comparison of external motion tracking systems for PET

list-mode reconstruction. *IEEE Nuclear Science Symposium Conference Record*, MIC15.S-191, 2011.

[115] O.V. Olesen, M. Sibomana, S.H Keller, F. Andersen, J.A. Jensen, S. Holm, C. Svarer, and L. Højgaard. Spatial resolution of the HRRT PET scanner using 3D-OSEM PSF reconstruction. *IEEE Nuclear Science Symposium Conference Record*, M13-225:3789–3790, 2009.

[116] O.V. Olesen, C. Svarer, M. Sibomana, S.H. Keller, S. Holm, J.A. Jensen, F. Andersen, and L. Højgaard. A movable phantom design for quantitative evaluation of motion correction studies on high resolution PET scanners. *IEEE Transactions on Nuclear Science*, 57(3):1116–1124, 2010.

[117] R.R. Paulsen, J.A. Bærentzen, and R. Larsen. Markov random field surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 16:636–646, 2010.

[118] R.R. Paulsen and R. Larsen. Anatomically plausible surface alignment and reconstruction. *Proceedings of Theory and Practice of Computer Graphics*, 249–254, 2010.

[119] A.N. Pedersen, S. Korreman, H. Nystrom, and L. Specht. Breathing adapted radiotherapy of breast cancer: reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold. *Radiotherapy and oncology*, 72(1):53–60, 2004.

[120] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.T. Chen. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *Journal of computer assisted tomography*, 13(1):20, 1989.

[121] F. Perruchot, A. Reilhac, C. Grova, A. C. Evans, and A. Dagher. Motion correction of multi-frame PET data. *IEEE Nuclear Science Symposium Conference Record*, 5:3186–3190, 2004.

[122] Y. Picard and C.J. Thompson. Digitized video subject positioning and surveillance system for PET. *IEEE Transactions on Nuclear Science*, 42(4):1024–1029, 1995.

[123] Y. Picard and C.J. Thompson. Motion correction of PET images using multiple acquisition frames. *IEEE Transactions on Medical Imaging*, 16(2):137–144, 1997.

[124] B. Planeta-Wilson, J. Yan, T. Mulnix, and R.E. Carson. Quantitative accuracy of HRRT list-mode reconstructions: effect of low statistics. *IEEE Nuclear Science Symposium Conference Record*, 5121–5124, 2008.

[125] E.B. Podgorsak. *Radiation Physics for Medical Physicists*. Springer, 2006.

[126] R. Poppe. Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, 108(1-2):4–18, 2007.

[127] F. Qiao, T. Pan, J.W. Clark, and O.R. Mawlawi. A motion-incorporated reconstruction method for gated PET studies. *Physics in Medicine and Biology*, 51:3769, 2006.

[128] F.H. Raab, E.B. Blood, T.O. Steiner, and H.R. Jones. Magnetic position and orientation tracking system. *IEEE Transactions on Aerospace and Electronic Systems*, 5(5):709–718, 1979.

[129] N. Raghunath, T.L. Faber, S. Suryanarayanan, and J.R. Votaw. Motion correction of PET brain images through deconvolution: II. practical implementation and algorithm optimization. *Physics in Medicine and Biology*, 54(3):813–829, 2009.

[130] A. Rahmim, P. Bloomfield, S. Houle, M. Lenox, C. Michel, K.R. Buckley, T.J. Ruth, and V. Sossi. Motion compensation in histogram-mode and list-mode EM reconstructions: beyond the event-driven approach. *IEEE Transactions on Nuclear Science*, 51(5):2588–2596, 2004.

[131] A. Rahmim, J.C. Cheng, K. Dinelle, M. Shilov, W.P. Segars, O.G. Rousset, B.M.W. Tsui, D.F. Wong, and V. Sossi. System matrix modelling of externally tracked motion. *Nuclear medicine communications*, 29(6):574–581, 2008.

[132] A. Rahmim, K. Dinelle, J.C. Cheng, M.A. Shilov, W.P. Segars, S.C. Lidstone, S. Blinder, O.G. Rousset, H. Vajihollahi, B. Tsui, D.F. Wong, and V. Sossi. Accurate event-driven motion compensation in high-resolution PET incorporating scattered and random events. *IEEE Transactions on Medical Imaging*, 27(8):1018–1033, 2008.

[133] A. Rahmim, J. Tang, and H. Zaidi. Four-dimensional (4D) image reconstruction strategies in dynamic PET: Beyond conventional independent frame reconstruction. *Medical physics*, 36:3654, 2009.

[134] M.K. Rasmussen. *Design and optimization of camera control on a clinical tracking system*. Bachelor thesis: Technical University of Denmark, Informatics:IMM–B.Eng.–2010–37, 2010.

[135] A.J. Reader, F.C. Sureau, C. Comtat, R. Trébossen, and I. Buvat. Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM. *Physics in medicine and biology*, 51:5455, 2006.

[136] M.H.T. Reinges, H.H. Nguyen, T. Krings, B.O. Hütter, V. Rohde, and J.M. Gilsbach. Course of brain shift during microsurgical resection of supratentorial cerebral lesions: limits of conventional neuronavigation. *Acta neurochirurgica*, 146(4):369–377, 2004.

[137] M. Reyes, G. Malandain, P.M. Koulibaly, M.A. Gonzalez-Ballester, and J. Darcourt. Model-based respiratory motion compensation for emission tomography image reconstruction. *Physics in Medicine and Biology*, 52:3579–3600, 2007.

[138] M. Rodriguez, J.S. Liow, S. Thada, M. Sibomana, S. Chelikani, T. Mulnix, C.A. Johnson, C. Michel, W.C. Barker, and R.E. Carson. Count-rate dependent component-based normalization for the HRRT. *IEEE Transactions on Nuclear Science*, 54(3):486–495, 2007.

[139] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. *Proceedings of International Conference in 3-D Digital Imaging and Modeling*, 145–152, 2001.

[140] U.E. Ruttimann, P.J. Andreason, and D. Rio. Head motion during positron emission tomography: Is it significant? *Psychiatry Research: Neuroimaging*, 61(1):43–51, 1995.

[141] G. Sansoni, S. Corini, S. Lazzari, R. Rodella, and F. Docchio. Three-dimensional imaging based on gray-code light projection: characterization of the measuring algorithm and development of a measuring system for industrial applications. *Applied Optics*, 36(19):4463–4472, 1997.

[142] H. Schreiber and J.H. Bruning. Phase shifting interferometry. *Optical Shop Testing*, 547–655, 2007.

[143] G.K. V. Schulthess. *Clinical Molecular Anatomic Imaging: PET, PET/CT and SPECT/CT*. Lippincott Williams and Wilkins, 2003.

[144] J. Schwider, R. Burow, K.E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel. Digital wave-front measuring interferometry: some systematic error sources. *Applied optics*, 22(21):3421–3432, 1983.

[145] S. Segobin, J. Matthews, D. Coope, S. Williams, M.Jones, R. Hinz, and K. Herholz. Motion detection from raw list-mode HRRT data and improved frame-by-frame realignment for motion correction. *Journal of Nuclear Medicine Meeting Abstracts*, 50:589, 2009.

[146] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, 2011.

[147] M. Sibomana, S.H. Keller, C. Svarer, O.V. Olesen, F. Andersen, S. Holm, and L. Højgaard. New attenuation correction for the HRRT using transmission scatter correction and total variation regularization. *IEEE Nuclear Science Symposium Conference Record*, 3284–3286, 2009.

[148] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23:208–219, 2004.

[149] D.L. Snyder. Parameter estimation for dynamic studies in emission-tomography systems having list-mode data. *IEEE Transactions on Nuclear Science*, 31(2):925–931, 1984.

[150] V. Srinivasan, HC Liu, and M. Halioua. Automated phase-measuring profilometry of 3-D diffuse objects. *Applied Optics*, 23(18):3105–3108, 1984.

[151] T.D.L. Steeves, J. Miyasaki, M. Zurowski, A.E. Lang, G. Pellecchia, T. Van Eimeren, P. Rusjan, S. Houle, and A.P. Strafella. Increased striatal dopamine release in parkinsonian patients with pathological gambling: a [11C] raclopride PET study. *Brain*, 132(5):1376, 2009.

[152] M.B. Stegmann and H.B.W. Larsson. Motion-compensation of cardiac perfusion MRI using a statistical texture ensemble. *Functional Imaging and Modeling of the Heart, Springer*, page 1009, 2003.

[153] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Medical Physics*, 24:25, 1997.

[154] F.C. Sureau, A.J. Reader, C. Comtat, C. Leroy, M.J. Ribeiro, and I. Buvat. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *Journal of Nuclear Medicine*, 49(6):1000, 2008.

[155] C. Svarer, K. Madsen, S.G. Hasselbalch, L.H. Pinborg, S. Haugbol, V.G. Frokjaer, S. Holm, O.B. Paulson, and G.M. Knudsen. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *Neuroimage*, 24(4):969–979, 2005.

[156] M. Takeda, H. Ina, and S. Kobayashi. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *Journal of the Optical Society of America*, 72(1):156–160, 1982.

[157] K. Thielemans. Scatter estimation and motion correction in PET. *IEEE Nuclear Science Symposium Conference Record*, 3:3, 2006.

[158] K. Thielemans, S. Mustafovic, and L. Schnorr. Image reconstruction of motion corrected sinograms. *IEEE Nuclear Science Symposium Conference Record*, 4:2401–2406, 2003.

[159] T.L. Troy and S.N. Thennadil. Optical properties of human skin in the near infrared wavelength range of 1000 to 2200 nm. *Journal of Biomedical Optics*, 6:167, 2001.

[160] C. Tsoumpas, J.E. Mackewn, P. Halsted, A.P. King, C. Buerger, J.J. Totman, T. Schaeffter, and P.K. Marsden. Simultaneous PET–MR acquisition and MR-derived motion fields for correction of non-rigid motion in PET. *Annals of nuclear medicine*, pages 1–6, 2010.

[161] C. Tsoumpas, F.E. Turkheimer, and K. Thielemans. A survey of approaches for direct parametric image reconstruction in emission tomography. *Medical physics*, 35:3963, 2008.

[162] J. Verhaeghe, P. Gravel, R. Mio, R. Fukasawa, P. Rosa-Neto, J.P. Soucy, C.J. Thompson, and A.J. Reader. Motion compensation for fully 4D PET reconstruction using PET superset data. *Physics in medicine and biology*, 55:4063–4082, 2010.

[163] Y. Wang. *Advances in Dynamic Radioative Scanning*. Charles C Thomas, Publisher, Springfield, Illinois, 1968.

[164] A.G. Weisenberger, S.S. Gleason, J. Goddard, B. Kross, S. Majewski, S.R. Meikle, M.J. Paulus, M. Pomper, V. Popov, M.F. Smith, et al. A restraint-free small animal SPECT imaging system with motion tracking. *IEEE Transactions on Nuclear Science*, 52(3):638–644, 2005.

[165] K. Wienhard, M. Schmand, and M.E. Casey. The ECAT HRRT: Performance and first clinical application of the new high resolution research tomograph. *IEEE Transactions on Nuclear Science*, 49(1):104–110, 2002.

[166] J. Wilm, O. Olesen, R. Paulsen, L. Højgaard, B. Roed, and R. Larsen. Real time surface registration for PET motion tracking. *Proceedings of the Scandinavian Conference on Image Analysis*, 6688:166–175, 2011.

[167] S.-K. Woo, H. Watabe, Y. Choi, K.M. Kim, C.C. Park, P.M. Bloomfield, and H. Iida. Sinogram-based motion correction of PET images using optical motion tracking system and list-mode data acquisition. *IEEE Transactions on Nuclear Science*, 51(3):782–788, 2004.

[168] R.P. Woods, S.R. Cherry, and J.C. Mazziotta. Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, 16(4):620–633, 1992.

[169] R.P. Woods, S.T. Grafton, C.J. Holmes, S.R. Cherry, and J.C. Mazziotta. Automated image registration: I. general methods and intrasubject, intramodality validation. *Journal of Computer Assisted Tomography*, 22(1):139, 1998.

[170] M. Zaitsev, C. Dold, G. Sakas, J. Hennig, and O. Speck. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *Neuroimage*, 31(3):1038–1050, 2006.

[171] S. Zhang and P. S. Huang. High-resolution, real-time three-dimensional shape measurement. *Optical Engineering*, 45(12):123601, 2006.

[172] S. Zhang and P. S Huang. Novel method for structured light system calibration. *Optical Engineering*, 45(8):83601–1 – 83601–8, 2006.

[173] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *International Conference on Computer Vision*, 1:666–673, 1999.

[174] V.W. Zhou, A.Z. Kyme, S.R. Meikle, and R. Fulton. An event-driven motion correction method for neurological PET studies of awake laboratory animals. *Molecular Imaging and Biology*, 10(6):315–324, 2008.